Data Analysis of Medical Images: CT, MRI, Phase Contrast X-ray and PET

Christensen, Anders Nymark

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

Data Analysis of Medical Images: CT, MRI, Phase Contrast X-ray and PET

Anders Nymark Christensen

Kongens Lyngby 2015
PhD-2015-386

Technical University of Denmark
Department of Applied Mathematics and Computer Science
Richard Petersens Plads, building 324,
2800 Kongens Lyngby, Denmark
Phone +45 4525 3351
compute@compute.dtu.dk
www.compute.dtu.dk PhD-2015-386 ISSN: 0909-3192

Summary (English)

Data analysis of medical images is an important and growing area, as systems for imaging becomes still more available and complex.

The goal of this thesis is to demonstrate solutions to data analysis problems in a cross disciplinary context. Further, to develop methods for analysis of new imaging modalities and to combine cross disciplinary knowledge from various fields to find new solutions to existing problems.

More specifically this thesis shows segmentation of images, classification and statistics used on a variety of quite different problems. Active Appearance models, Chan-Vese and graph-cut has been used, as well as a variety of statistical tools centred on the General Linear Model.

The point of departure for the thesis is the NanoGuide project, in which gel based x-ray markers for use in radiotherapy has been developed. Two different types of gels has been analysed using segmentation of micro-CT images followed by a statistical analysis of homogeneity, contrast, degradation, and other qualities. By combining knowledge from the different professions in the project, a new application for one of the developed gels - in-vivo dosimetry in radiotherapy - has been studied.

Analysis of differences between groups and of correlations between brain regions and cognitive tests in alzheimers patients is another contribution. Segmentation of fat in abdominal MRI-scans has also been studied and a robust algorithm based on graph-cut is presented.

A relatively new modality – phase-contrast x-ray and dark-field – has shown promise for diagnosis of a variety of diseases in the lungs. A classification algorithm for differentiation of healthy, emphysematous and fibrotic lung tissue on pixel level is presented.

Resumé

Dataanalyse af medicinske billeder er et vigtigt og voksende område, da systemer til billeddannelse bliver stadig mere udbredte og komplekse.

Målet for denne afhandling er at demonstrere løsninger på dataanalyseproblemer af medicinske billeder i en tværfaglig sammenhæng. Endvidere, at udvikle metoder til analyse af billeder fra nye billeddannede modaliteter og kombinere tværfaglig viden fra forskellige felter til at finde nye løsninger på eksisterende problemer.

Mere specifikt viser denne afhandling segmentering af billeder, klasifikation og statistik anvendt på en række - ret forskellige - problemer. Active Apperance Models, Chan-Vese og graph-cut er blevet brugt, og en række statistiske metoder, centreret omkring den Generelle Lineære Model er blevet anvendt.

Udgangspunktet for afhandlingen er NanoGuideprojektet, i hvilket gelbaserede røntgenmarkører til brug i stråleterapi er blevet udviklet. To forskellige typer af geler er blevet analyseret ved hjælp af segmentering i mikro-CT billeder og en efterfølgende statistisk analyse af homogenitet, kontrast, nedbrydelse og andre egenskaber. Ved at kombinere viden fra de forskellige faggrupper i projektet, er en ny anvendelse - in-vivo dosimtry i stråleterapi - for de udviklede geler blevet undersøgt.

Analyse af forskel mellem grupper og korrelation mellem segmenterede hjerneregioner og kognitive test i alzheimerspatienter udgør et andet bidrag. Segmentering af fedt i abdominale MR-scaninnger er også blevet undersøgt, og en robust algoritme baseret på graph-cut præsenteres.

En relativt ny modalitet – fasekontrast-røntgen og mørkefelt – har vist sig lovende til diagnoisticering af en række sygdomme i lungerne. En klassifikationsalgoritme til at differentiere sundt, emfysemt og fibrøst lungevæv på pixelniveau præsenteres.

Preface

This thesis was prepared at the department of Applied Mathematics and Computer Science at the Technical University of Denmark in fulfilment of the requirements for acquiring an PhD in Applied Mathematics with an emphasis on Mathematical Image Analysis. The thesis was part of the NanoGuide project and was prepared with funding from the Danish Council for Strategic Research. Professor Knut Conradsen started as co-supervisor and ended as the main supervisor and vice versa for professor Rasmus Larsen.

The thesis deals with data analysis of medical images, namely CT, MRI, Phase Contrast X-ray and PET.

The thesis consists of a part with the general methodology that was applied and a part summarising the scientific contributions of the thesis.

Lyngby, 01-October-2015

Anders Nymark Christensen

Contributions

Papers included in this thesis

A Jølck, R. I., Rydhög, J. S., Christensen, A. N., Hansen, A. E., Bruun, L. M., Schaarup-Jensen, H., von Wenck, A. S., Børresen, B., Kristensen, A. T., Clausen, M. H., Kjaer, A., Conradsen, K., Larsen, R., af Rosenschöld, P. M., and Andresen, T. L. (2015). *Injectable Colloidal Gold for Use in Intrafractional 2D Image-Guided Radiation Therapy.* Advanced Healthcare Materials, 4(6):856-863.

B Christensen, A. N. (2013). Segmentation and estimation of the degradation rate over 14 weeks of two formulations of the x-SAIB gel in mice. Technical report

C Christensen, A.N., Rydhög, J. S., af Rosenschöld, P. M., Andresen, T. L., Holm, S., Conradsen, K. Jølck, R. I. (2015). *Injectable Silver Nanogels: In Vivo Dosimetry for External Beam Radiotherapy using Positron Emission Tomography.* Under submission.

D Larsen, C. T., Frederiksen, K. S., Hasselbalch, S. G., Christensen, A. N., Høgh, P., Wermuth, L., Lolk, A., Andersen, B. B., Siebner, H.R., Waldemar, G., Garde, E. (2015). *Effect of moderate-to-high intensity aerobic exercise on hippocampus and cortical regions in patients with mild to moderate Alzheimer's disease.* Under submission.

E Christensen, A. N., Larsen, C. T., Petersen, M. B., Conradsen, K., Andersen, V. D. (2015). *Software-Package for Automated Bias-Field Correction and Quantifcation of Abdominal Fat using MRI with Preliminary Evaluation in Overweight Subjects.* Technical Report.

F Christensen, A.N., Yaroshenko, A., Hellbach, K., Willer, K., Gromann, L., Yildirim, A. Ö., Conradsen, K., Pfeiffer, F. (2015). *Lung disease classification in mice using x-ray phase-contrast and dark-field imaging.* Under submission.

Other work by the author

- Abstract:
  Christensen, A. N. Reichkendler, M., Auerback, P., Larsen, R., Nielsen, H., Ploug, T., Stallknecht, B., Hø jgaard1, L., and Holm1, S. (2012). *Non-invasive Estimation of Metabolic Uptake Rate of Glucose using F18-FDG PET and Linear Transformation of Outputs.* In 25th congress of the European Association of Nuclear Medicine.

- Abstract:
  Reichkendler, M., Auerbach, P., Rosenkilde, M., Christensen, A. N., Holm, S., Sjödin, A., Kjaer, A., Ploug, T., Højgaard, L., and Stallknecht, B. (2013). *Differential effects of aerobic exercise on insulin-stimulated glucose uptake in skeletal muscle and adipose tissue examined with PET/CT.* In Journal of Nuclear Medicine, page 54 (Supplement 2):255.

- Article:
  Reichkendler, M. H., Auerbach, P., Rosenkilde, M., Christensen, a. N., Holm, S., Petersen, M. B., Lagerberg, a., Larsson, H. B. W., Rostrup, E., Mosbech, T. H., Sjödin, a., Kjaer, a., Ploug, T., Hoejgaard, L., and Stallknecht, B. (2013). *Exercise training favors increased insulin-stimulated glucose uptake in skeletal muscle in contrast to adipose tissue: a randomized study using FDG PET imaging.* American journal of physiology. Endocrinology and metabolism, 305(4):E496-506.

- Article:
  Christensen, A. N., Reichkendler, M. H., Larsen, R., Auerbach, P., Højgaard, L., Nielsen, H. B., Ploug, T., Stallknecht, B., and Holm, S. r. (2014). *Calibrated image-derived input functions for the determination of the metabolic uptake rate of glucose with [18F]-FDG PET.* Nuclear medicine communications, pages 8-14.

- Commentary:
  Christensen, A. N. (2014). *Commentary to 'Application of calibrated image-derived input function to a clinical protocol'.* Nuclear Medicine Communications, 35(11):1189-1190.

- Book:
  Andersen, F., Antoch, G., Belohlavek, O., Beyer, T., Boanova, L., Bock-

isch, A., Boellaard, R., Christensen, A., Dahlbom, M., Delis, H., Desousa, A., Herzog, H., Holm, S., Jø dal, L., Kabickova, E., Kalemis, A., Kinahan, P., Loft, A., Löfgren, J., Machado, M., Mawlawi, O., Muzi, M., Palm, S., Pan, T., Poli, G., Ribeiro, M., J., F., Robilotta, C., Schäfers, K., Teles Garcez, A., and Townsend, D. (2014). *IAEA HUMAN HEALTH SERIES: PET / CT Atlas on Quality Control and Image Artefacts.* Number 27.

- Abstract:
  Henning, W., Andersen, J., Christensen, A., Greisen, G., Højgaard, L., and Law, I. (2014). *Validation of a new technique to estimate regional cerebral blood flow in piglets using fluorescent microspheres.* J. Nucl. Med., 55(sup. 1):1804.

- Article:
  Henning, W., Andersen, J. B., Højgaard, L., Greisen, G., Law, I., Thorseth, A., and Christensen, A. N. (2015). *A new self-made digital slide scanner and microscope for imaging and quantifcation of fluorescent microspheres.* Journal of Biomedical Graphics and Computing. 5, 2, p. 33-39.

- Abstract:
  Tinggaard, J., Hagen, C., Mouritsen, A., Mieritz, M., Wohlfahrt-Veje, C. Fallentin, E., Larsen, R., Christensen, A., Jensen, R., Juula, A., and Maina, K. (2015). *Abdominal fat Distribution Measured by Magnetic Resonance Imaging in 197 Children Aged 10-15 Years – Correlation to Anthropometry and Dual X-Ray Absorptiometry.* In ESPE Abstracts, pages 84 P-2-330.

Acknowledgements

I would like to thank my supervisors professor Knut Conradsen and professor Rasmus Larsen. Their guidance and helpfulness have been an invaluable contribution to this work. My colleagues over the three years have been very pleasant company. In particular Jacob Schack Vestergaard, Stine Harder, Anders Lindbo Larsen, Oula Puonti, Martin Kjer, Gudmundur Einarson, and Chrstian Thode Larsen deserves my thankfulness in connection with this thesis.

The many people involved in the NanoGuide project has been one of the driving forces. I would like to thank Jonas Scherman Rydhög and Rasmus Irming Jølck for a fruitful collaboration.

My external stay at Technische Universität München, would not have been possible without professor Franz Pfeiffer, and my stay would have been a lot less rewarding - both intellectually and socially - if not for Andre Yaroshenko, Lukas Gromann and Konstantin Willer. In the same context I should mention Hildur Einarsdottir for generously allowing me free use of her code.

Last, but certainly not least, I would like to thank my girlfriend Sidsel, for putting up with my constantly decreasing presence at home and my, paradoxically, increasingly messy desk in our living room.

Contents

Summary (English) i

Resumé iii

Preface v

Contributions vii

Acknowledgements xi

1 Introduction 1
1.1 NanoGuide project . . . 2
1.1.1 Motivating examples . . . 3
1.2 Brains, Abdominal Fat and Mice . . . 5
1.3 Reading Guidelines . . . 7

I Methodology 9

2 Imaging, Radiotherapy and Related Physics 11
2.1 Computed tomography . . . 11
2.1.1 Cone Beam Computed Tomography . . . 12
2.2 Phase Contrast X-ray . . . 13
2.3 Magnetic Resonance Imaging . . . 15
2.4 Positron emission tomography . . . 18
2.4.1 Reconstruction algorithms . . . 19
2.5 Radiotherapy . . . 20
2.5.1 Targeting . . . 21
2.6 Photo-nuclear reactions . . . 22

3 Image Analysis **25**
3.1 Segmentation . . . 25
3.1.1 Active Appearance Models . . . 25
3.1.2 Chan-Vese . . . 26
3.1.3 Graph-Cut . . . 28
3.2 Image Filters . . . 30
3.2.1 Image smoothing . . . 30
3.2.2 Texture filters . . . 31
3.2.3 Feature extraction . . . 31
3.3 Histogram equalisation . . . 33
3.4 Bias Correction . . . 33

4 Statistics **35**
4.1 General Linear Model . . . 35
4.1.1 Analysis of variance . . . 37
4.1.2 t-test . . . 38
4.1.3 Welch's t-test . . . 38
4.1.4 Multivariate analysis of variance . . . 39
4.1.5 Hotellings T^2 . . . 40
4.1.6 Estimation in the General Linear Model . . . 41
4.2 Repeated Measurements Model . . . 43
4.2.1 Covariance structure in the errors . . . 45
4.3 Assumption tests . . . 45
4.3.1 Influence statistics and residuals . . . 46
4.3.2 Normality tests . . . 49
4.3.3 Variance tests . . . 49
4.4 Model Selection . . . 50
4.4.1 R^2 and adjusted R^2 . . . 51
4.4.2 Information Criteria . . . 51
4.5 Non-parametric methods . . . 52
4.5.1 Mann-Whitney-U . . . 52
4.5.2 Multivariate rank tests . . . 53
4.6 Variance stabilising transforms . . . 54
4.6.1 Box-Cox . . . 54
4.7 Correlation tests . . . 55
4.8 Performance measures . . . 56
4.9 Classification . . . 57
4.9.1 Naive Bayes . . . 57
4.9.2 Discriminant Analysis . . . 58
4.9.3 Logistic Regression . . . 58
4.9.4 Random forest . . . 59
4.9.5 Support Vector Machines . . . 59
4.9.6 Sparse methods . . . 61

II Scientific Contributions 63

5 NanoGuide project: Paper A, B, and C 65
5.1 Paper A . . . 65
5.1.1 Methods and Results . . . 66
5.1.2 Contributions . . . 66
5.2 Paper B . . . 67
5.2.1 Methods and Results . . . 67
5.2.2 Contributions . . . 70
5.3 Paper C . . . 70
5.3.1 Methods and Results . . . 70
5.3.2 Contributions . . . 71

6 Paper D 73
6.1 Methods and Results . . . 73
6.2 Contributions . . . 74

7 Paper E 75
7.1 Methods and Results . . . 75
7.2 Contributions . . . 76

8 Paper F 77
8.1 Methods and Results . . . 77
8.2 Contributions . . . 78

9 Conclusion 81

III Included Papers 83

A Injectable Colloidal Gold for Use in Intrafractional 2D Image-Guided Radiation Therapy 85

B Segmentation and estimation of the degradation rate over 14 weeks of two formulations of the x-SAIB gel in mice 103

C Injectable Silver Nanogels: In Vivo Dosimetry for External Beam Radiotherapy using Positron Emission Tomography 137

D Effect of moderate-to-high intensity aerobic exercise on hippocampus and cortical regions in patients with mild to moderate Alzheimer's disease 157

E Software-Package for Automated Bias-Field Correction and Quantifcation of Abdominal Fat using MRI with Preliminary Evaluation in Overweight Subjects **183**

F Lung disease classification in mice using x-ray phase-contrast and dark-field imaging **207**

Bibliography **219**

Chapter 1

Introduction

Data analysis of medical images are an important and growing field. With the advent and wide spread use of diverse systems for medical imaging, we face an increasing demand for automation and quantitative analysis. Both to save valuable time in the clinic, to extract information not apparent to a human observer, and to interpret results rigorously to avoid spurious conclusions. Further by combining physical knowledge of the various systems, we may develop new applications to solve current problems.

The objectives of this thesis are:

- To engineer solutions to data analysis problems of medical images in cross disciplinary settings
- To develop methodologies for new medical imaging techniques
- To combine data analysis with physical knowledge of imaging modalities to achieve new solutions to exiting problems

1.1 NanoGuide project

Cancer is diagnosed in more than 14.1 million people each year and 8.2 million die from it [1]. With the ageing populations and better diagnostics this number is increasing.

For the treatment more than 50 % of cancer patients receive radiotherapy. Radiotherapy uses ionising radiation to kill the cancerous cells. This treatment is highly dependent on advanced imaging techniques such as CT, PET and MRI for the location of the cancer. A major challenge is the movement of both tumours and normal tissue during the radiotherapy.

These problems can potentially be alleviated by image-guiding during the treatment. However, the tumours are often difficult to visualise and track with the imaging apparatus available in connection with the treatment. To increase the image contrast, solid markers of e.g. gold are implanted in or around the tumours. This requires large diameter needles and adverse side-effects are relatively common.

The goal of the NanoGuide project was thus to develop new markers and tracers for use in the treatment of cancer, based on nano-technology, that did not have the short-comings of the existing markers and were better than existing tracers.

This thesis is part of NanoGuide project and the first three contributions also come from this project. The project is cross-disciplinary in nature and has contributions from DTU Nanotech, DTU Chemistry, DTU Nutech and DTU Compute through yours truly and my supervisors. Also contributing are Copenhagen University Hospital, Rigshospitalet, Gentofte and Bispebjerg and the industry in the form of the companies Nanovi and GE Healthcare. In this setting the demands of the project took precedence, with regard to the tasks and scientific work that were pursued.

It was originally envisioned, that a systemic marker for intra-venous injection was going to be developed. The nano-particles in this marker would then circulate with the bloodstream and bind to any cancerous tissue, marking them for radiotherapy. In that regard my task would have been analysis of CT- and PET-images investigating the biodistribution of the tracer. Unfortunately, this part of the project met unforeseen challenges and has not yet come to fruition.

A second marker for local injection through a thin flexible needle was pursued concurrently with the systemic marker. This marker - based on gold-

[1] `http://www.cancerresearchuk.org/health-professional/cancer-statistics/worldwide-cancer`

nanoparticles (paper A) and later iodine (paper B) - was much more successful. It was originally envisioned that shape-statistics of the marker would be hugely relevant. Both for the characterisation of the marker in terms of stability and for its use in radiotherapy. This turned out to be of very little practical importance, and the effort was put into general characterisation of the gels with regard to stability and visibility.

The iodine based gel-marker is in the final stages of regulatory approval, and has been tested in-vivo in both mice, pigs and patients. It was decided in the project not to publish results along the way to avoid damaging the novelty value. As such all the results are being collected into a single high-impact article, describing the full process from bench-to-bedside. In this thesis a technical report has been included (paper B), exemplifying the work I have contributed. Many other tasks have been performed with regard to statistical analysis of results and general modelling, both for regulatory approval and for other smaller experiments.

As often seen in science within engineering the work was thus problem driven and had to be adjusted to the requirements of the NanoGuide-project. While little new methodology has been developed, existing techniques has been applied in new settings and for new purposes.

1.1.1 Motivating examples

In external beam radiotherapy we rely on accurate and precise targeting of the tumour. Even with precise delineation and alignment of the patient, breathing and bowel movements will lead to possible errors. More over, with conventional 2D x-ray imaging the contrast is too low to track the tumour on-line while treating. One of the goals in the NanoGuide project was to develop a nanoparticle-hydrogel marker for local administration (fluid that gels upon injection) with good contrast. Thus allowing for online tracking of the tumor while treating. This has been achieved, see paper A, where gold nano-particles were used to achieve the necessary contrast. For the analysis we needed to quantify the inhomogeneity in the gels. The gels were placed in mice and imaged on a micro-CT scanner. The analysis was made more complicated by the fact, that gels with different contrast levels were compared.

A second even more promising gel is in the final stages of regulatory approval, see paper B. It uses iodine instead of gold, and has very good characteristics with regard to stability and contrast. An important subtask for me, has been statistical analysis of results for regulatory approval of the developed markers. In the paper we had a situation with multiple measurements on the same mice.

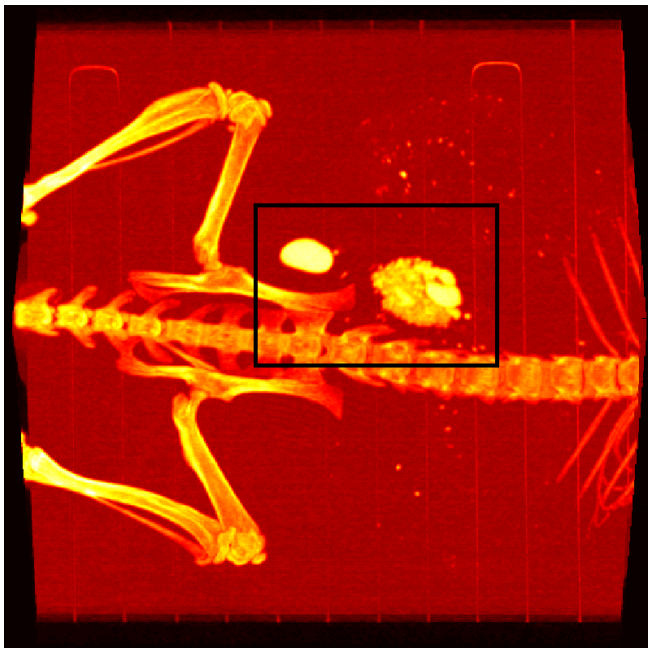


Figure 1.1: Maximum Intensity Projection of two gels in a mouse, paper B

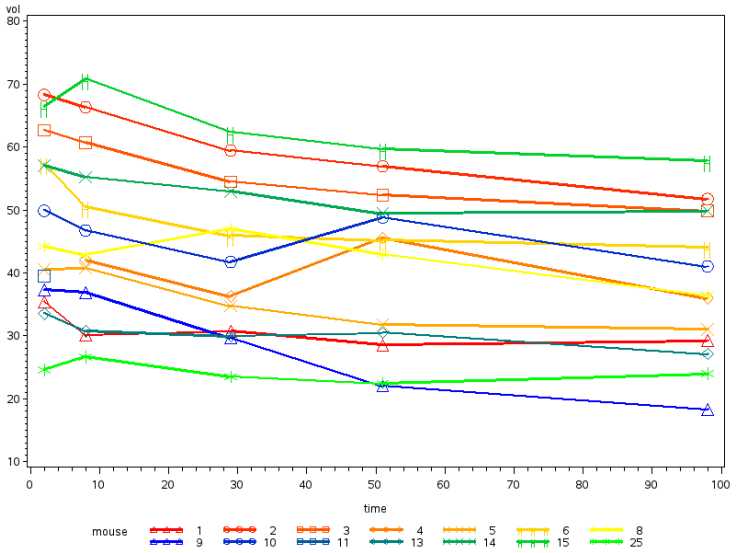


Figure 1.2: Volume data over time from mice in paper B

We first needed to segment the gel, see fig. 1.1, and could then extract volume information, see fig. 1.2. Here the main interest was the degradation rate of different formulations of the gel. A Repeated Measurements Model was used to quantify the results.

Through this collaboration another possible application for the gels developed was envisioned, namely as an in-vivo dosimeter, see paper C.

By combining knowledge about nuclear physics with knowledge about problems in radiotherapy and the capabilities of the different imaging modalities available, we can reach new solutions. One is to measure the dose delivered through radiotherapy in-vivo, by activating silver nano-particles embedded in a gel through a (γ,n)-reaction. By measuring the resulting activity level in a PET scanner, we can estimate the dose. The activation of the silver gels are shown in fig. 1.3.

1.2 Brains, Abdominal Fat and Mice

Although this thesis took its point of departure in the NanoGuide-project, the basic methodology, image-analysis and statistical analysis could be readily applied to other areas. This exemplifies the general use of the approach.

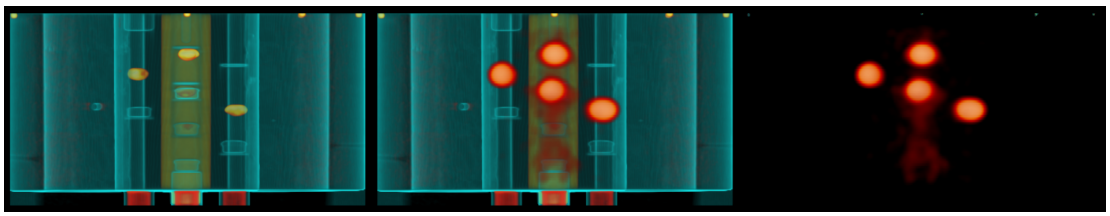
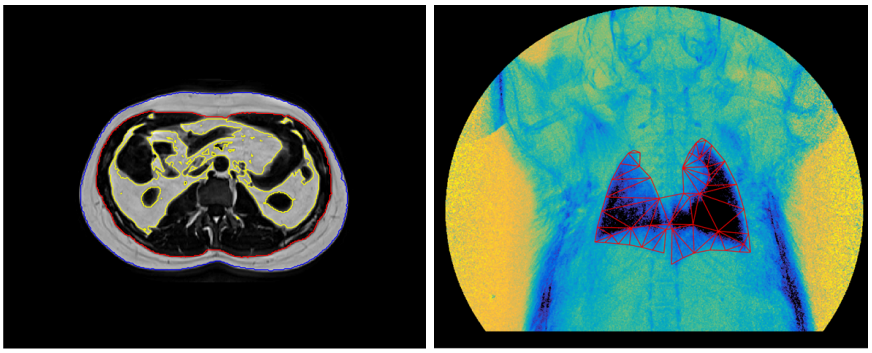


Figure 1.3: Maximum intensity projections. From left to right CT, PET/CT, and PET. The silver gels are clearly activated by the photon irradiation used in radiotherapy. See paper C.



(a) Segmentation of abdominal fat, paper E

(b) Lung segmentation in a mouse using a dark-field image, paper F

Figure 1.4: Different segmentation problems

In paper D the effect of exercise on alzheimers disease was investigated through MRI-scans of the brain. To quantify changes between two groups a Hotellings T^2 test was performed. As several outliers were present in the data, the results were further confirmed by a non-parametric test. In the same paper it was investigated whether there was any correlation between cognitive test scores and the changes in specific brain regions. A statistical test for this special situation was employed.

In paper E we had a segmentation problem: How to quantify the amount of abdominal fat from a MRI-scan. Faced with a segmentation problem, the standard approach outside the image analysis community is nearly always manual segmentation. This is slow, tedious work, and suffers from both inter- and intra-observer variance. By making these segmentations automatic, we can both reduce the variance stemming from the observer and drastically reduce the required time for analysis. See fig. 1.4 (a) for an example of the segmentation.

In paper F we were presented with the challenge of classifying diseases in mice, based on absorption, dark-field and differential phase images. First the lungs needed to be segmented, see fig. 1.4 (b), to extract the relevant parts. Fig. 1.5 illustrates a classification on the basis of only two of the images available: absorption and dark-field.

1.3 Reading Guidelines

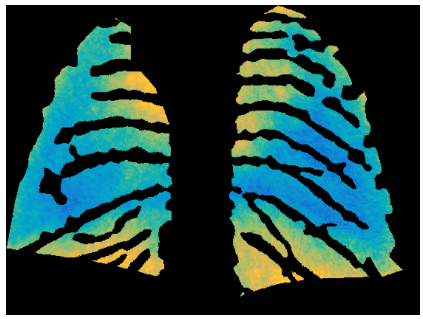
Part I serves as a common reference and introduction to the methodologies used in this thesis. This part assumes a basic knowledge of physics, image analysis and statistics.

Part II summarises the contributions and puts them in context.

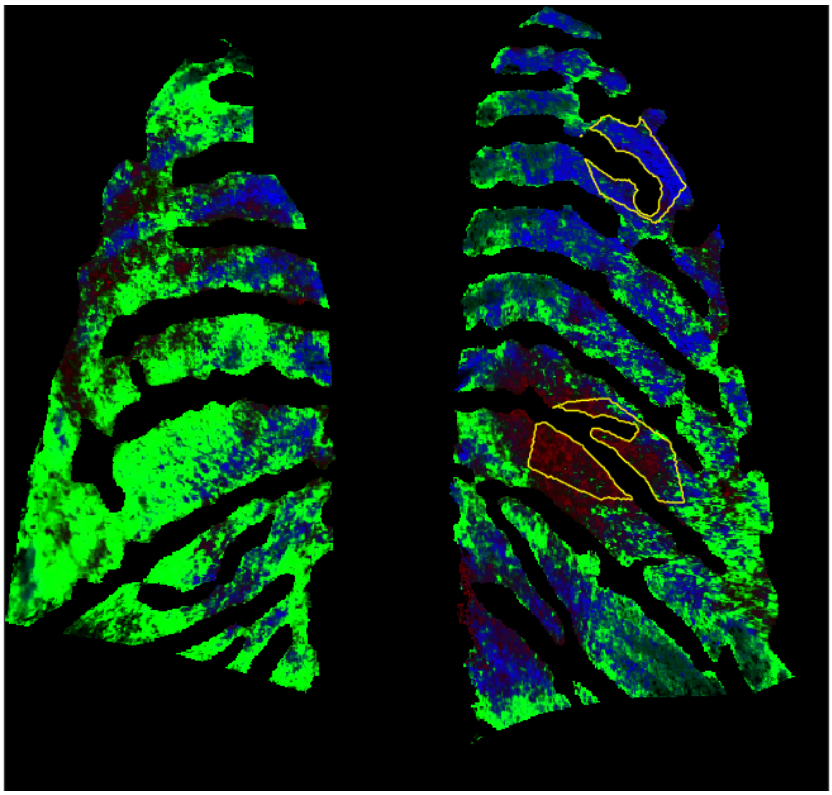
The contributions are included in part III.

The thesis can either be read from the beginning, or one can start with part II and refer back to part I to clarify details in the applied methodology.

(a) Absorption image



(b) Dark-field image



(c) Classification

Figure 1.5: Illustration of classification. Green is healthy, blue is emphysema and red is fibrotic. The marked regions are classified as fibrotic by a radiologist.

Part I

Methodology

Chapter 2

Imaging, Radiotherapy and Related Physics

This chapter is meant as a short overview of the imaging modalities used in this thesis, as well as the basic principles behind radiotherapy and some of the related physics. References are given if the reader want a more in-depth treatment.

2.1 Computed tomography

Computed tomography (CT) uses an X-ray tube and detector panel mounted around an axis where the object or subject to be imaged are placed. By rotating the tube and detector together, images are taken at different angles. Given enough of these angular projections a 3D image can be reconstructed using Filtered Back Projection (FBP), see section 2.4.1.1. While having a rather poor contrast in soft tissue, CT has excellent contrast between soft tissue and bone.

A CT scan is quantitative and the unit of the images are normally given in Hounsfield Units (HU) defined by:

$$HU = 1000 \cdot \left(\frac{\mu}{\mu_{\text{water}}} - 1 \right) \tag{2.1}$$

where μ is the linear attenuation coefficient. Table 2.1 shows the HU values in various tissues.

Table 2.1: HU values in different substances and organs [Jørgensen, 2005]

Substance	HU
Air	-1000
Lung	-900 to -500
Fat	-100 to -85
Water	0
Kidney	20 to 40
White matter	20 to 40
Grey matter	35 to 55
Blood	50 to 60
Muscle	40 to 50
Liver	60 to 70
Spongious bone	30 to 230
Compact bone	250 to 3000

Both resolution, energy of the x-rays (kV) and the amount of photons used (mAs) can be varied. This gives a trade-off between image quality and the dose delivered to the subject. The typical dose can range from about 1 mSv to more than 10 mSv depending on much of the body is covered by the scan and the quality of the scan [Elliott, 2011].

See e.g. [Bushberg et al., 2011] for a more in-depth treatment of CT. Examples of uses can be seen in paper A where it was used to quantify the contrast level, volume, and homogeneity in a gold-based NanoGel. In paper B it was used to extract volumetric information about a iodine-based NanoGel, and in paper C it was used for location of gels, dose planning and attenuation correction.

2.1.1 Cone Beam Computed Tomography

A Cone Beam Computed Tomography (CBCT) integrated with a radiotherapy treatment apparatus, was first introduced in 2002 [Jaffray et al., 2002]. The principle is the same as for CT, but a cone beam geometry is used. It allows for in-situ control of the patient placement, by comparing a planning CT with the CBCT images.

The geometry with a cone of X-rays hitting a flat-panel detector, does not allow for the resolution of quantitative units (HU). The geometry also induces a number of artefacts [Schulze et al., 2011]. Iterative reconstruction algorithms

that can correct for some of these artefacts exists. They are, however, rather slow for clinical use. Instead a variation of the FBP called the Feldkamp algorithm is used [Feldkamp et al., 1984]. Even with these drawbacks, CBCT is still a valuable tool for quality control in radiotherapy.

Examples of uses can be seen in papers A and C.

2.2 Phase Contrast X-ray

CT and normal x-ray images rely only on the absorption of the x-rays. In phase contrast imaging – as the name implies – the phase is also used.

Several techniques exists, but only grating based imaging (GBI) will be treated here. See [Bech, 2009, Nesterets and Wilkins, 2008] for an overview of the different techniques. The review here for GBI is based on [Bech, 2009, Jensen, 2010].

The complex refractive index is given by:

$$n = 1 - \delta + i\beta \tag{2.2}$$

where δ is the normal refractive index determining the phase shift and β is the attenuation, see fig. 2.1.

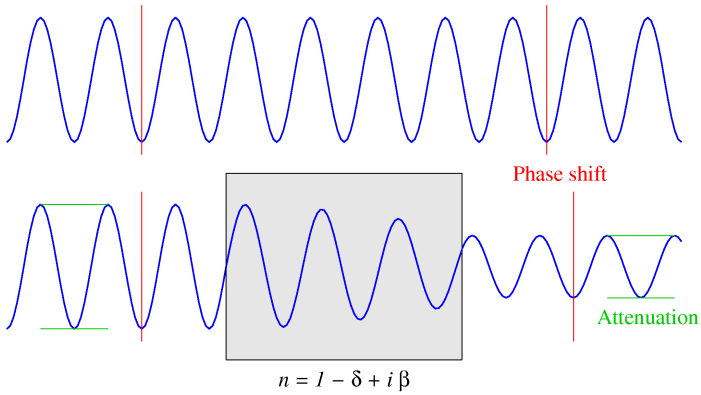


Figure 2.1: Illustration of the complex refractive index. We see the resulting phase shift and attenuation by passing through a medium. From [Bech, 2009]

We want an imaging technique that, in addition to measuring β, the attenuation, measures the phase shift δ. The grating based imaging can do that, and more. The basic mechanism is given by fig. 2.2.

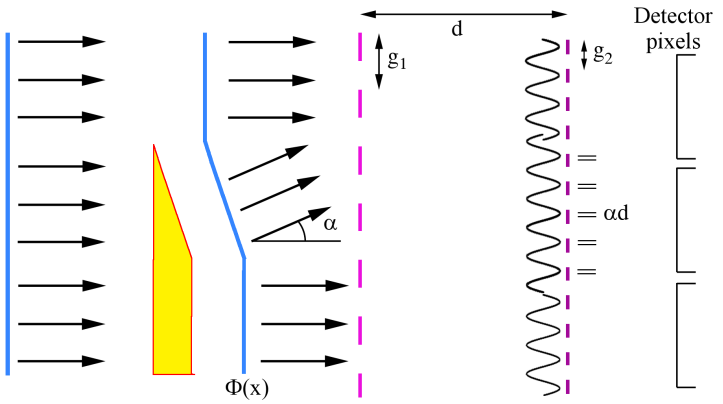


Figure 2.2: Illustration of the two gratings that make detection of phaseshift possible. From [Bech, 2009]

By introducing the $g1$ grating, we get a diffraction pattern. However, to resolve that we would need very high resolution in our detector. By introducing the second grating $g2$ we can sample the wavefront, by shifting the $g2$ grating, thus measuring the phase shift at several positions and resolve it.

However, this setup needs coherent x-rays, but it is possible to insert a grating array that acts as many thin slits, enabling the use of standard x-ray tubes. In that case we have a setup like the one seen in fig. 2.3.

The positions and the characteristics of the gratings needs to be matched to the x-ray energy, and the distance from the source to $g1$.

When shifting the $g2$ grating the signal can be written as:

$$I_{p_x,p_y}(x_g) = a_0(p_x,p_y) + \sum_{m=1}^{\infty} a_m(p_x,p_y)\cos\left(\frac{2\pi m x_g}{g_2} + \phi_m(p_x,p_y)\right) \qquad (2.3)$$

where x_g is the shifting, I_{p_x,p_y} is the measured intensity in pixel x, y.

Instead of fitting the intensities directly to this equation to resolve, a_0, a_m and, ϕ_m, we can simply take the Fourier transform. In addition to the sample scan ^{s}, a reference scan ^{r} is acquired, to correct for inhomogeneities.

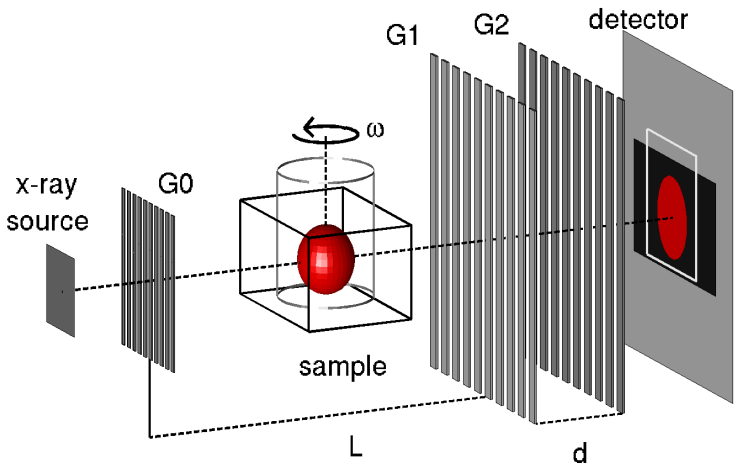


Figure 2.3: GBI setup with incoherent x-ray source. From [Bech, 2009]

We can now recover the absorption or transmission image by:

$$a_0 = \frac{a_0^s}{a_0^r} \tag{2.4}$$

the differential phase is given by:

$$\phi_1 = \phi_1^s - \phi_1^r \tag{2.5}$$

and the visibility – normally known as darkfield – as:

$$V = \frac{a_0^s}{a_0^r} \sum_{m \text{ odd}} \frac{a_m^s}{a_m^r} \tag{2.6}$$

examples of the three images can be seen in fig. 2.4. Examples of uses can be seen in paper F, where the modality was used to classify between diseased and healthy lung in a murine model.

2.3 Magnetic Resonance Imaging

Magnetic Resonance Imaging (MRI) is a versatile imaging technique that utilises nuclear magnetic resonance (NMR). While the images are not quantitative in nature like CT, it can be used to get quantitative results, especially with regards to volume. The following review is based on [Nishimura, 2010, Hanson, 2009].

(a) Absorption image **(b)** Dark-field image **(c)** Diff. phase image

Figure 2.4: The three image types obtainable using grating based imaging. The images are of the same mouse.

In relation to imaging of humans, it is the hydrogen content of the body that is used, as it has non-zero spin. Associated with the spin is the magnetic moment $\mathbf{M}$. Given a huge magnet, that generates a magnetic field $\mathbf{B}$, the moment and field will reach an equilibrium over time and point in the same direction.

If we now perturb the moment by a radio frequency (RF) pulse we get:

$$\frac{d\mathbf{M}}{dt} = \mathbf{M} \times \mathbf{B} \tag{2.7}$$

Remembering the right hand rule, we see that $\mathbf{M}$ will precess around $\mathbf{B}$. The frequency is given by: $f = \frac{\gamma}{2\pi} B$ Hz, where γ is the gyromagnetic ratio and the frequency is known as the Larmor frequency.

To do imaging we use RF pulses to introduce specific flips in the magnetic moment. Further we use gradient coils to introduce a gradient in the magnetic field $\mathbf{B}$ across the subject. As the frequency is dependant on $\mathbf{B}$ f will vary across the subject and a roll pattern will develop. The returned signal is now dependent on how much this roll pattern is similar to the subject. In effect we sample the subject in the Fourier domain or k-space, as it is called within MRI. An illustration can be seen in fig. 2.5.

This gives us a trade-off between scan time and resolution. If we only cover the central parts of k-space, we will get a blurred image. If we want to cover k-space more extensively it increases the scan time.

Many different scans can be performed that weigh the subject differently. The three most used sequences are T1, T2 and proton-density. But other sequences are also used, i.e. Dixon [Ma, 2008, Dixon, 1984] which is very fast and gives two images: one of the water content and one of the fat content, but at the cost of poor resolution. If not constrained by speed T1 with water suppression seems the best for imaging fat.

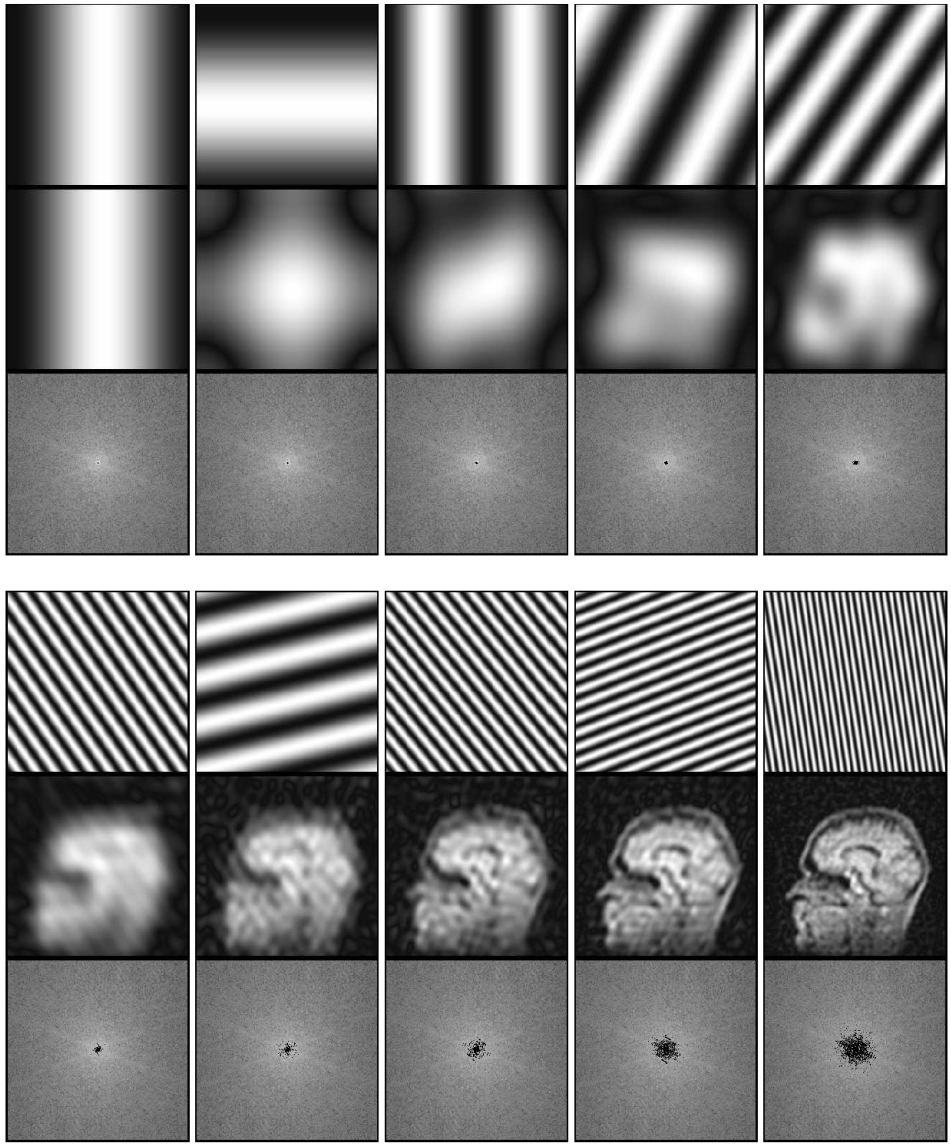


Figure 2.5: Row 1: More and more detailed roll patterns are added. Row 2: The resulting image as more patterns are added. Row 3: The parts of k-space that have been sampled. From [Hanson, 2009]

As the scans are relatively long, a common artefact is subject movement. Another common artefact is a bias-field. This is due to inhomogeneities in the

magnetic field **B**, and leads to varying intensity across the images.

Examples of uses can be seen in paper D where it was used to extract volumetric information about the brain in alzheimers patients and in paper E where the modality is used for quantification of abdominal fat.

2.4 Positron emission tomography

Positron emission tomography (PET) detects the γ-photons that are emitted following the annihilation of a positron from a β^+ decay and an electron. The following is based on [Knöß, 2004].

A ring of detectors are placed around the subject, and when two photons are detected at approximately the same time (within approx. 6 ns [Knöß, 2004]), we know that a decay has happened on the line between the two detectors: The line of response (LOR). The newer PET scanners can even resolve the time at a level detailed enough (picosecond level!) to give information about where on the line the decay has occurred.

The source of the β^+-decays are a radioactive tracer injected into the subject. Some of the more commonly used isotopes are ^{18}F, ^{11}C and ^{15}O.

Several sources of errors must be considered and corrected for PET scans to yield quantitative results. Three of the most important are:

- **Attenuation:** The subject scanned attenuates the γ-photons. If no attenuation correction is made, the signal for superficial parts will be overestimated compared to deeper lying parts. To correct for the attenuation most newer systems use a CT scan, integrated into the PET scanner [Kinahan et al., 2003]. Some systems use a radioactive line or point source that are rotated around the subject to estimate it.
- **Scatter:** The γ-photons may be scattered on their way to the detectors. This means that the decay has not taken place on the LOR. If not corrected, this will lead to a blurring of the entire image. Newer systems use model based, monte-carlo simulations incorporating information from the CT scan that was also used for attenuation correction [Siemens, 2009].
- **Decay:** The tracer injected decays while the scan is performed. Typically the scan is decay corrected to the start of the scan. The decay formula is:

$$A(t) = A_0 e^{-\lambda t} \tag{2.8}$$

  where A_0 is the initial activity and λ the decay constant.

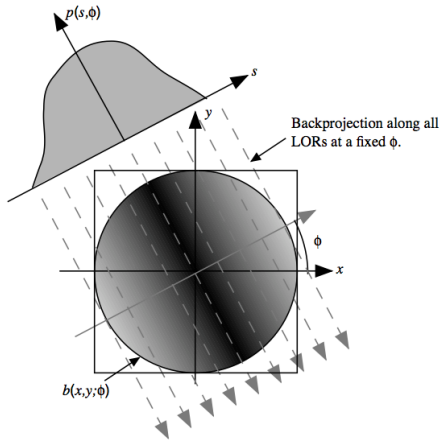


Figure 2.6: The sinogram is formed from the different angles ϕ and positions s. Copyright © 2013, Milan Zvolský [1]

An example of uses can be seen in paper C, for the measurement of radioactive silver used for in-vivo dosimetry.

2.4.1 Reconstruction algorithms

The detected events are collected into a sinogram. The sinogram has angles along one axis and position perpendicular to the given angle along the other axis , see fig. 2.6. For reconstructing the sinograms into images, two main groups of algorithms exist. The filtered backprojection (FBP) and the iterative reconstructions.

2.4.1.1 Filtered Backprojection

The FBP is generally the most reliable algorithm, but also yields more noisy images than the iterative methods.

If we just back-projected these data, we would get a heavy blurring in the middle. Instead we take the Fourier transformation, apply a filter to avoid the

[1] `http://www.desy.de/~garutti/LECTURES/BioMedical/Lecture7_ImageReconstruction.pdf`

blurring, and finally apply the inverse Fourier transformation.

$$f(x,y) = \mathcal{F}^{-1}\{h(\omega)\mathcal{F}(v_x, v_y)\} \tag{2.9}$$

where $\mathcal{F}$ is the Fourier transform, h is the filter, and v_x, v_y are polar coordinates.

2.4.1.2 iterative reconstruction

The iterative reconstruction method derive from the Ordered Subset Expectation Maximisation method [Hudson and Larkin, 1994]. They have been developed to incorporate the points-spread function and other corrections, to increase the resolution and reduce noise, see e.g. [Comtat et al., 2008]. The forward model is given by:

$$\mathbf{y} = \mathbf{A}\boldsymbol{\lambda} + \mathbf{b} \tag{2.10}$$

where $\mathbf{y}$ is the sinogram, $\boldsymbol{\lambda}$ is the image we want to estimate, and $\mathbf{b}$ is the noise and scatter. $\mathbf{A}$ is a system matrix that contains both geometric information, but also information about the blurring effects. Because of this the backwards model can get quite complicated.

A specific problem with iterative methods are, that they tend to overestimate low activities [van Velden et al., 2009, Boellaard et al., 2001]. For uses where we want quantitative results the FBP should thus be preferred.

2.5 Radiotherapy

Radiotherapy aims at destroying cancerous tissue by delivering doses of ionising radiation, while minimising side effects on healthy tissue. The ionising radiation induces DNA damage, and as cancerous tissue has a higher rate of cell division, is is more susceptible to the damage.

While particle based radiation has seen growing use, it is still very specialised and expensive. Most commonly used is high-energy photons, which are generated by Bremsstrahlung from linear accelerators. By using a multi-leaf collimator, the x-ray beam can be precisely shaped, see fig. 2.7.

A CT scan is used for the dose plan. Radiographers will outline the tumor or tissue to be targeted, and medical physicists will make a dose-plan using

[2] `http://filecache.drivetheweb.com/mr5mr_varian/154577/MLC+-+gold.jpg`
`http://filecache.drivetheweb.com/mr5mr_varian/154631/clinac+dose+shaping.jpg`

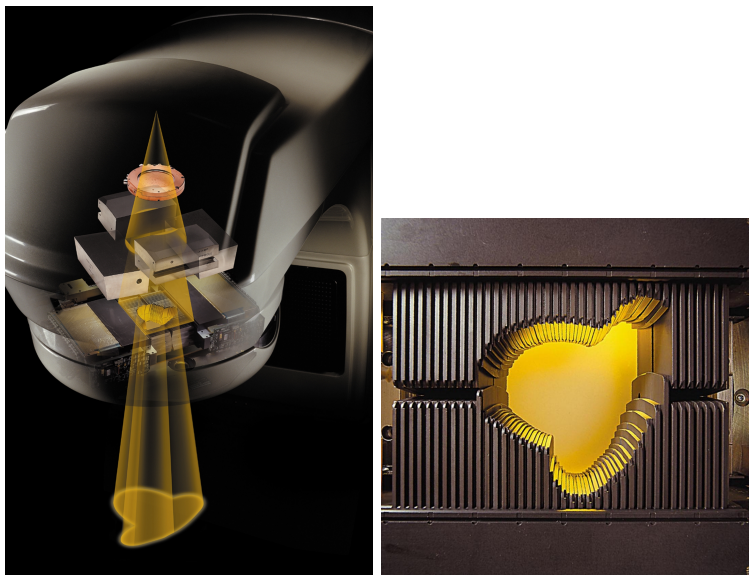


Figure 2.7: Beam forming using a multi-leaf collimator. Copyright © Varian Medical [2]

advanced monte-carlo based models. By irradiating from multiple angles at varying intensities and shapes, a sufficient dose can be delivered to the target tissue, while minimising dose to vulnerable tissue. An example can be seen in fig. 2.8.

Examples of uses can be seen in paper A and C.

2.5.1 Targeting

A very important subproblem in radiotherapy is accurate targeting. If we assume the radiographers have contoured the cancerous tissue correctly, we need to target it precisely.

CBCT is a big help here, as it can be used to verify correct patient placement in the accelerator. However, for tumours in thorax or abdomen we have displacements due to breathing and bowel movements. Unfortunately the 2D x-ray

[3] `http://filecache.drivetheweb.com/mr5mr_varian/154488/Base+of+Tongue+RapidArc+Axial.jpg`

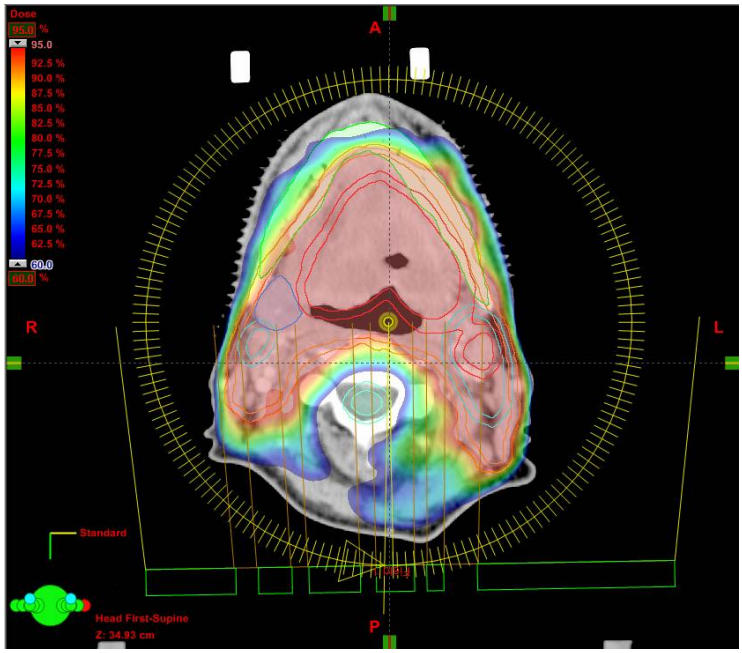


Figure 2.8: Dose plan for a mouth cancer. We note how the spinal cord is spared for excissive radiation, and receives a low dose by using the entire arc of angles. Copyright © Varian Medical [3]

images that we can record online while treating cannot be used to distinguish the tumour, as the contrast is too low. It has been one of the aims of the NanoGuide-project to rectify this by injectable x-ray markers.

2.6 Photo-nuclear reactions

In the context of radiotherapy the (γ,n)-reaction has interesting applications. A high energy photon can knock a neutron out of the nucleus, producing a new isotope. If the produced isotope decays via β^+, it can be imaged in a PET scanner and potentially be used for dosimetry.

The required energy for the reaction to take place varies from isotope to isotope. Of interest are the isotopes naturally occurring in the body and silver, which has a high cross-section and produces a high fraction of β^+ decay. Other materials have been investigated, but silver was the most promising [Hansen et al., 2008].

However, the energies required for the isotopes naturally occurring in the body – ^{12}C: 20.7 MeV, ^{14}N: 17.5 MeV, ^{16}O: 15.7 MeV – are higher than normally used in radiotherapy and the cross sections are small [Fuller, 1985]. Further even if activated, we have no ground truth of the amount of the isotope present, which makes quantification very difficult.

In contrast silver can be placed in the subject in a known amount. Natural silver consists of the isotopes ^{106}Ag and ^{108}Ag, and the energy threshold is only 9.5 MeV [Diven and Almy, 1950]. Their decay is outlined in fig. 2.9.

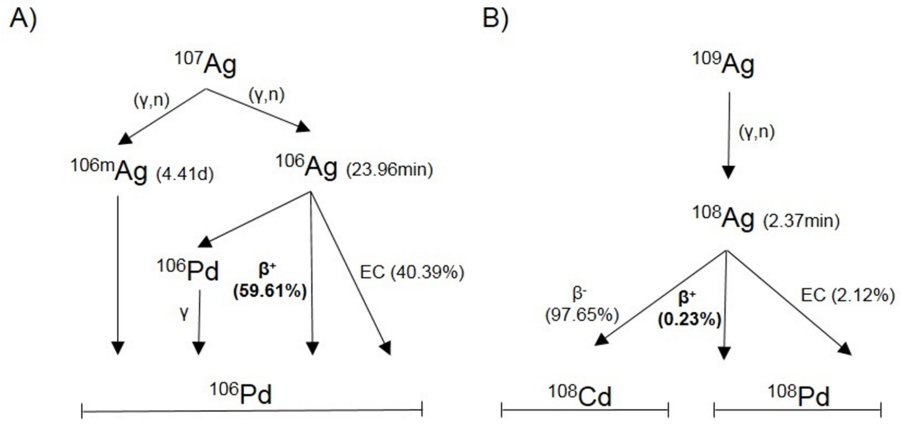


Figure 2.9: Decay of ^{106}Ag and ^{108}Ag. From [Hansen et al., 2008]

An example of uses can be seen in paper C, where the reaction is used for in-vivo dosimetry.

Chapter 3

Image Analysis

This chapter will give a brief overview of the image analysis techniques used in this thesis. Further details can be found in the references.

3.1 Segmentation

Segmentation is a basic problem in image analysis, and depending on the context, many algorithms are available, ranging from a simple threshold to complex models. Below we will review some techniques of importance to this thesis.

3.1.1 Active Appearance Models

The Active Appearance Model (AAM) has seen widespread applications since its introduction in 1998 [Cootes et al., 1998]. The idea is to fit a model or shape to an image by using not only the shape but also the appearance of the image. This makes it a more robust alternative to active shape models [Cootes et al., 1995, Cootes and Taylor, 2001].

We start with a training set of images with annotated landmarks of a shape and all in the same coordinate system. Let each shape be represented by a vector $\mathbf{x}$. Then applying a principal component analysis (PCA) on the training set, we can describe each shape as:

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{P}_s \mathbf{b}_s \tag{3.1}$$

where $\bar{\mathbf{x}}$ is the overall mean shape, $\mathbf{P}_s$ is the principal components, in this context often referred to as modes of variation, and $\mathbf{b}_s$ is the shape parameters.

We can now warp each shape $\mathbf{x}$ to the mean shape $\bar{\mathbf{x}}$, thus shape normalising them. Given the shape normalised images we extract the convex hull of the shape and the appearance - or image intensity - can now be modelled equivalently:

$$\mathbf{g} = \bar{\mathbf{g}} + \mathbf{P}_g \mathbf{b}_g \tag{3.2}$$

where $\mathbf{g}$ is the mean and intensity normalised image patches.

By including scaling, translation, rotation, and intensity scaling offset in the parameters, the shape of a new image can be estimated. To do it efficiently all the parameters are systematically perturbed from their empirical optimal values and averaged over the dataset. Having this change matrix precomputed speeds up the estimation significantly. In addition a multi-scale approach is used: Starting at a coarse level and then refining.

An example of the model used to segment the lungs in a mouse can be seen in fig. 3.1. The model was trained by Hildur Einarsdottir [Einarsdottir et al., 2015] and used in paper F.

3.1.2 Chan-Vese

Chan and Vese proposed to attack the segmentation problem through a level-set approach [Chan and Vese, 2001]. Let u denote the image we are trying to segment and let ϕ be a level-set function describing a curve and thus partitioning our image. Further, let H be the Heaviside function:

$$H(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \tag{3.3}$$

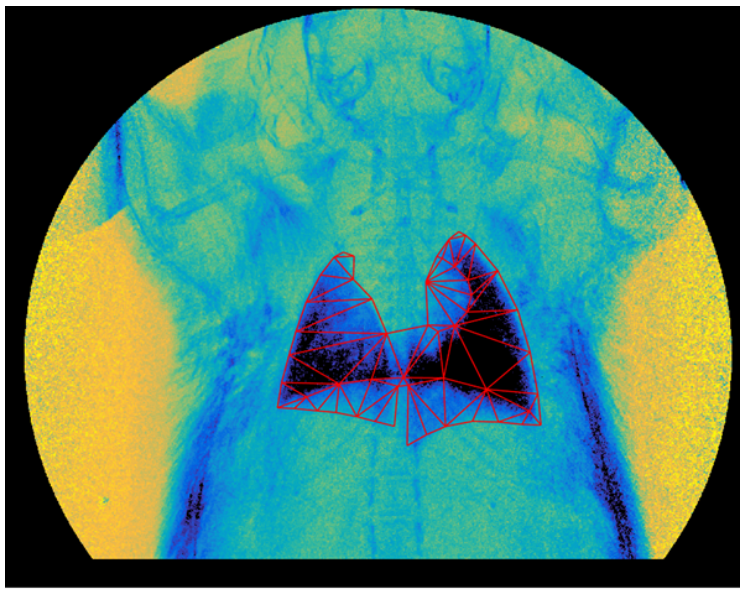


Figure 3.1: AAM used for lung segmentation in a mouse

The method then seeks to minimise an energy functional given by:

$$\begin{aligned} F(\phi) =& \mu \int_{\Omega} |\nabla H(\phi(x,y))| \; dx \; dy \\ &+ \nu \int_{\Omega} H(\phi(x,y)) \; dx \; dy \\ &+ \lambda_1 \int_{\Omega} |u(x,y) - c_1|^2 H(\phi(x,y)) \; dx \; dy \\ &+ \lambda_2 \int_{\Omega} |u(x,y) - c_2|^2 (1 - H(\phi(x,y))) dx \; dy \end{aligned} \tag{3.4}$$

where Ω describes the entire image. The first term is the length of the curve weighted by μ. The second term is the area enclosed by the curve weighted by ν. The two last terms are the homogeneity of the area inside and outside the curve, where c_1, c_2 are the mean value of the areas, weighted by λ_1 λ_2.

Often ν is set to 0, and the two last terms concatenated into one with equal weighting. In addition some degree of smoothing is often applied to counter the effects of noise. This is the case in the implementation[1] used to segment gels in mice in paper B.

[1] `http://www.mathworks.com/matlabcentral/fileexchange/24998-2d-3d-image-segmentation-toolbox`

3.1.3 Graph-Cut

Graph-cut allows for 'optimal' segmentations as graph-cut algorithms can be used to find the optimum cut in a graph. The problem is how to formulate ones segmentation problem as a graph.

Boykov and Kolmogorov published an efficient algorithm in 2004 [Boykov and Kolmogorov, 2004] for graph-cuts in vision, although graph-cuts had seen application before, see e.g. [Ishikawa, 2003, Ebastien, 1999].

We denote a directed graph $\mathcal{G} = \langle \mathcal{V}, \mathcal{E} \rangle$, which consists of nodes with individual costs $\mathcal{V}$ and directed edges also with individual costs $\mathcal{E}$. Further, the graph has a source t and a sink s. The task is now to find the minimum s/t-cut, that partitions the graph into two subsets. This is done by the min-cut / max-flow property. Looking at the graph as a set of pipes with different capacity, it seems obvious that the pipes saturated and thus determining the max-flow across the graph, will be the same pipes that lead to a minimum cut.

To utilise it in segmentation we need to formulate the image as a graph, see fig. 3.2.

This is the most basic formulation. Often the node and edge weight are given by the Gibbs model [Geman and Geman, 1984] leading to Markov Random Fields [Greig et al., 1989], but can be extended in several different ways [Boykov and Kolmogorov, 2003].

In this thesis a method for multi-layer segmentation proposed by Li et al. [Li et al., 2006] was used.

Given a 3D image we can view it as a matrix $\mathcal{I}(\mathbf{x}, \mathbf{y}, \mathbf{z})$. If the orientation is as in fig. 3.3 (a) the surface will only intersect each $(\mathbf{x}, \mathbf{y})$ column of voxels once. To ensure connectivity we have intracolumn edges E^a given by:

$$E^a = \{\langle V(x, y, z), V(x, y, z-1)\rangle | z > 0\} \tag{3.5}$$

To incorporate smoothness constraints Δ_x and Δ_y we add intercolumn edges given by:

$$E^r = \begin{cases} \{\langle V(x, \mathbf{y}, z),\ V(x+1, \mathbf{y}, \max(0, z-\Delta_x))\rangle \mid x \in \{0, \ldots, X-2\}, z \in \mathbf{z}\} \cup \\ \{\langle V(x, \mathbf{y}, z),\ V(x-1, \mathbf{y}, \max(0, z-\Delta_x))\rangle \mid x \in \{1, \ldots, X-1\}, z \in \mathbf{z}\} \cup \\ \{\langle V(\mathbf{x}, y, z),\ V(\mathbf{x}, y+1, \max(0, z-\Delta_y))\rangle \mid y \in \{0, \ldots, Y-2\}, z \in \mathbf{z}\} \cup \\ \{\langle V(\mathbf{x}, y, z),\ V(\mathbf{x}, y-1, \max(0, z-\Delta_y))\rangle \mid y \in \{1, \ldots, Y-1\}, z \in \mathbf{z}\} \end{cases} \tag{3.6}$$

The intercolumn edges guarantee, that if voxel $\mathcal{I}(x, y, z)$ is on the the surface $\mathcal{N}$, then the neighboring voxels on $\mathcal{N}$ in the x-direction, cannot be lower than than

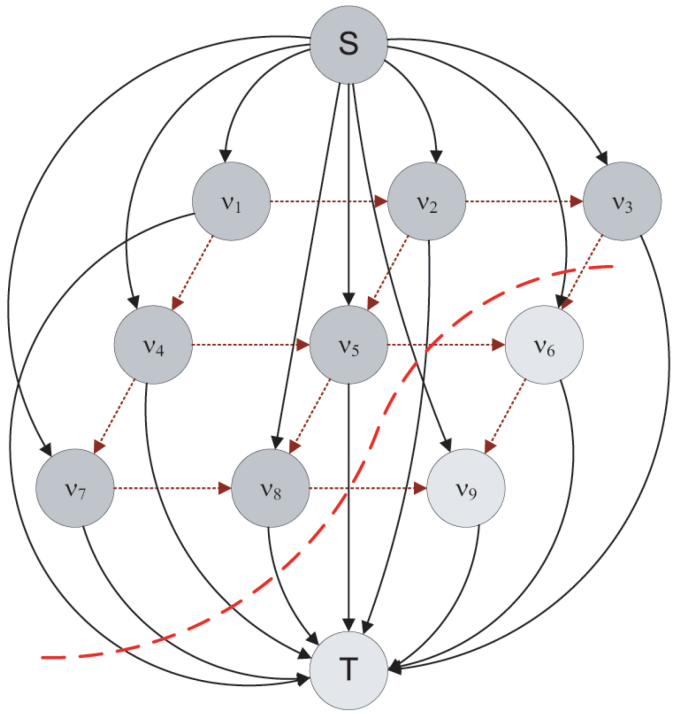


Figure 3.2: Typical grid-like graph for segmentation in images. While shown as uni-directional here, the image edges (red arrows) would typical be bidirectional. The cut is the dashed line. From: *Graph cuts via l1 norm minimization, Bhusnurmath et al, IEEE transactions on pattern analysis and machine, 2008*, Copyright © 2008, IEEE

$\mathcal{I}(x, y, \max(0, z - \Delta_x))$ and similar for the y-direction. The setup is illustrated in fig. 3.3 (b).

This can be generalised to 4D, when we have multiple surfaces. The constraints now ensure that we have minimum and maximum distance between the surfaces. The node weights or energies are set depending on what we want to segment. If it is layers in a 3D volume, it could e.g. be the inverse of the gradient. Often the surface we want to segment is folded or circular. The above method can still be used after an unrolling.

In paper E the method was used for segmentation of subcutaneous fat and in F it was used to identify the spine of mice on x-ray images.

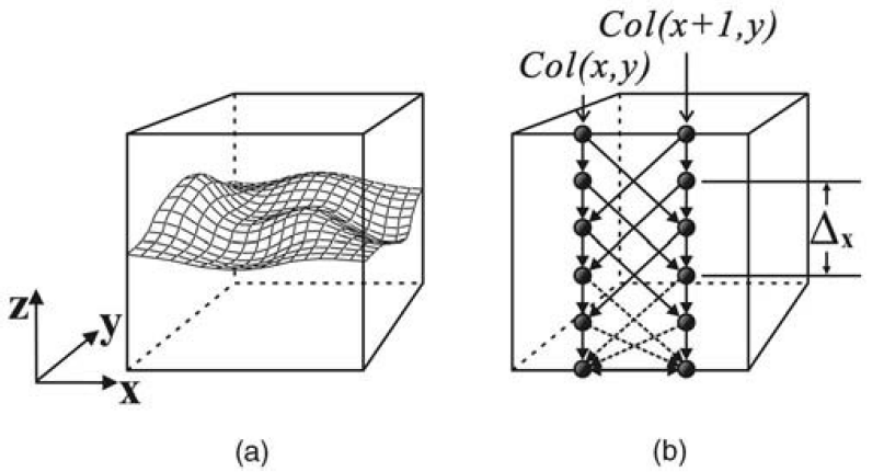


Figure 3.3: (a) Surface orientation. (b) Two adjacent columns with edges. [Li et al., 2006], Copyright © 2006, IEEE

3.2 Image Filters

An image can be treated as a matrix, and we can perform the same operations on an image as on a matrix, even though a lot of them might not be sensible. We will look at some of the more common filters, that have also been applied in the contributions of this thesis.

By kernel we mean a mask, that we center on each pixel or element in the image, calculate the value from element-wise multiplication of the kernel and image and then replace the pixel we have centred on with the resulting value.

3.2.1 Image smoothing

An image can be smoothed, i.e. noise removed or reduced. The most common – and some of the most simple – methods are Gaussian and median smoothing.

We start by choosing a 2D-Gaussian kernel:

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{3.7}$$

with a given σ. This function is discretised and made into a $n \times n$ matrix. This size will depend on the size of σ. A larger standard deviation needs a larger matrix.

The median smoothing is a non-linear filter, and cannot be formulated as a kernel filter. Within the kernel area the median value is calculated and that value returned to the centre pixel. The degree of smoothing is now directly dependant on the size of the matrix.

For both types of kernel we have assumed symmetry. This needs not be the case. If we want a heavier smoothing in one direction, we can make the kernel larger in that direction and vice versa. The methods saw use in paper E and F.

3.2.2 Texture filters

To describe a local feature like texture, we can also use filters.

The range filter is given by a kernel of size $n \times m$. Within that kernel, the range is calculated and the values returned. The same for the standard deviation filter except the standard deviation is returned. The entropy filter is a bit more complicated. It calculates the entropy $\mathcal{E}$ in the kernel given by:

$$\mathcal{E} = -\sum_{\text{kernel}} p \cdot \log_2(p) \tag{3.8}$$

where p is the histogram approximation to the probability density function from the kernel region.

The texture filters were used in paper F to extract information from the differential phase image.

3.2.3 Feature extraction

We can also use filters to extracts features from the image, i.e. edges or ridges.

A very versatile filter is the Laplacian of Gaussian (LoG) filter. The kernel is given by:

$$LoG(x, y) = -\frac{1}{4\pi\sigma^4}\left[1 - \frac{x^2 + y^2}{2\sigma^2}\right] e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{3.9}$$

which is simply the Laplacian of the Gaussian kernel. The filter can be used to identify blobs, holes, ridges and valleys, depending on the sign.

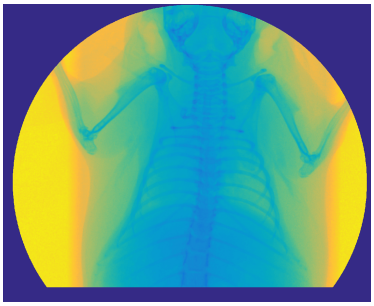
The gradient can be used to detect edges. A simple kernel is given by:

$$G_x = \begin{bmatrix} -\frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} \tag{3.10}$$

$$G_y = \begin{bmatrix} -\frac{1}{2} \\ 0 \\ \frac{1}{2} \end{bmatrix} \tag{3.11}$$

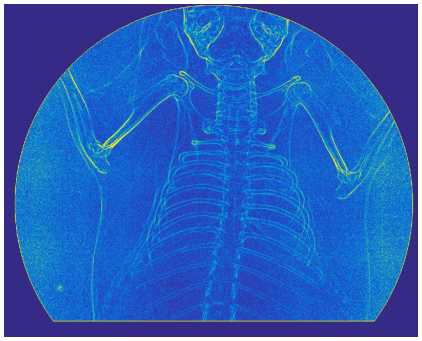
which is simply the central difference in the x and y direction.

An example of the two filters is given in fig. 3.4. In addition to paper F, the filers were also used extensively in paper E for the estimation of energies for segmentation.



(a) Absorption image of mouse

(b) LoG



(c) Gradient

Figure 3.4: Example of gradient and LoG filter used in paper F. (b) The LoG filter is with a kernel of 25×25 and with a σ of 8. (c) Sum of the absolute value of the gradients in the x- and y-directions.

3.3 Histogram equalisation

When posed with an image with poor contrast, we can enhance it by histogram equilisation. The basic idea is that by making the histogram of the image more uniformly distributed, we use more colours and the contrast is increased.

Since the contrast is not globally dependant, but is local feature, it also makes sense to do this in local areas of the image instead of globally.

One method is the Contrast Limited Adaptive Histogram Equilisation (CLAHE) [Zuidervel, 1994]. The image is partitioned into 8x8 regions, and a histogram equilisation is performed in each region. Since this alone often leads to noise in homogeneous regions and many bins 'wasted' on homogeneous regions, a clipping factor is introduced. This clipping factor determines when the histogram should be 'clipped' and bins with values above are thus redistributed equally. See fig. 3.5. The effect can be seen in fig. 3.6. Both the number of regions and the clipping factor can be varied.

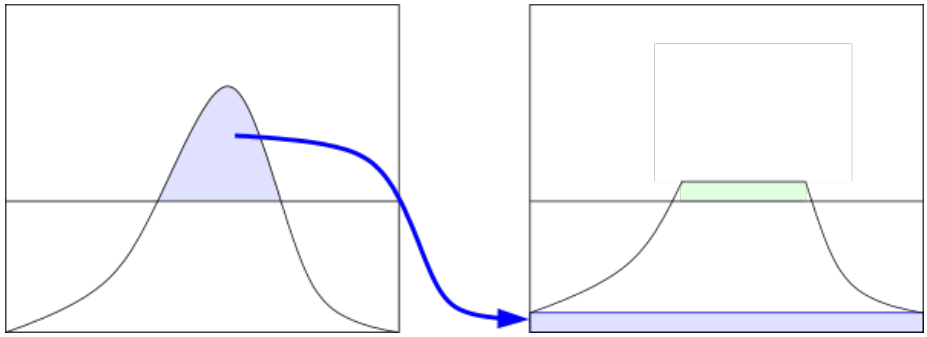
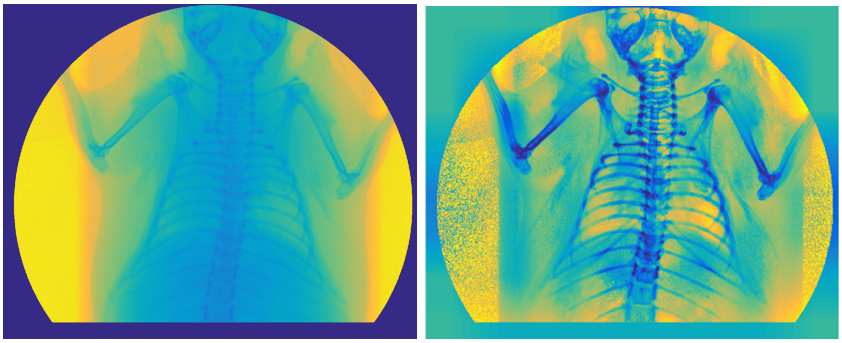


Figure 3.5: Clipping and redistribution procedure in CLAHE [2]

3.4 Bias Correction

Bias field correction has a very important role in MRI, but also has applications in other areas e.g. to correct for unequal illumination in vision systems. Using a generative model [Larsen, 2015] we can estimate and model the bias-field by a sum of basis functions.

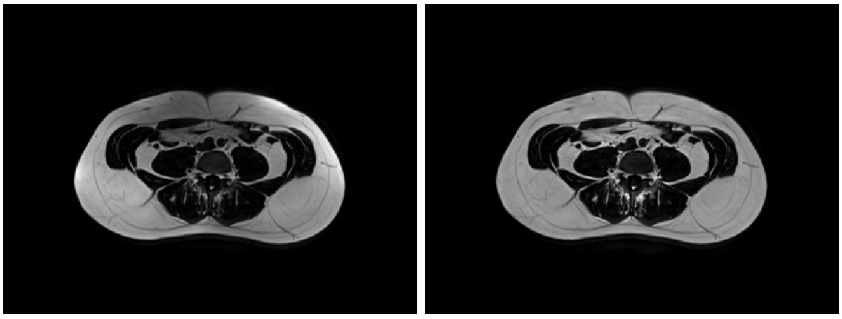
[2] https://upload.wikimedia.org/wikipedia/commons/thumb/5/5f/Clahe-redist.svg/600px-Clahe-redist.svg.png



(a) Absorption image of mouse **(b)** CLAHE

Figure 3.6: Example of the effect of contrast equalisation. Used in paper F for identification of ribs

The image is assumed to consist of a number of Gaussian classes, which can be set depending on the scan. By tuning this and other hyperparameters such as a smoothing distance controlling how flexible the bias-field is and a regularisation parameter punishing curvature of the bias-field, we can achieve very good corrections. For an example see fig. 3.7.



(a) MRI with bias-field **(b)** MRI corrected for bias-field

Figure 3.7: Example of a MRI before and after bias-correction. Used in paper E

Chapter 4

Statistics

This chapter is meant to give an overview and a common reference for the statistical methods used in this work. Although it is very brief it has more of a survey character then the preceding chapters. Further details may be found in the references.

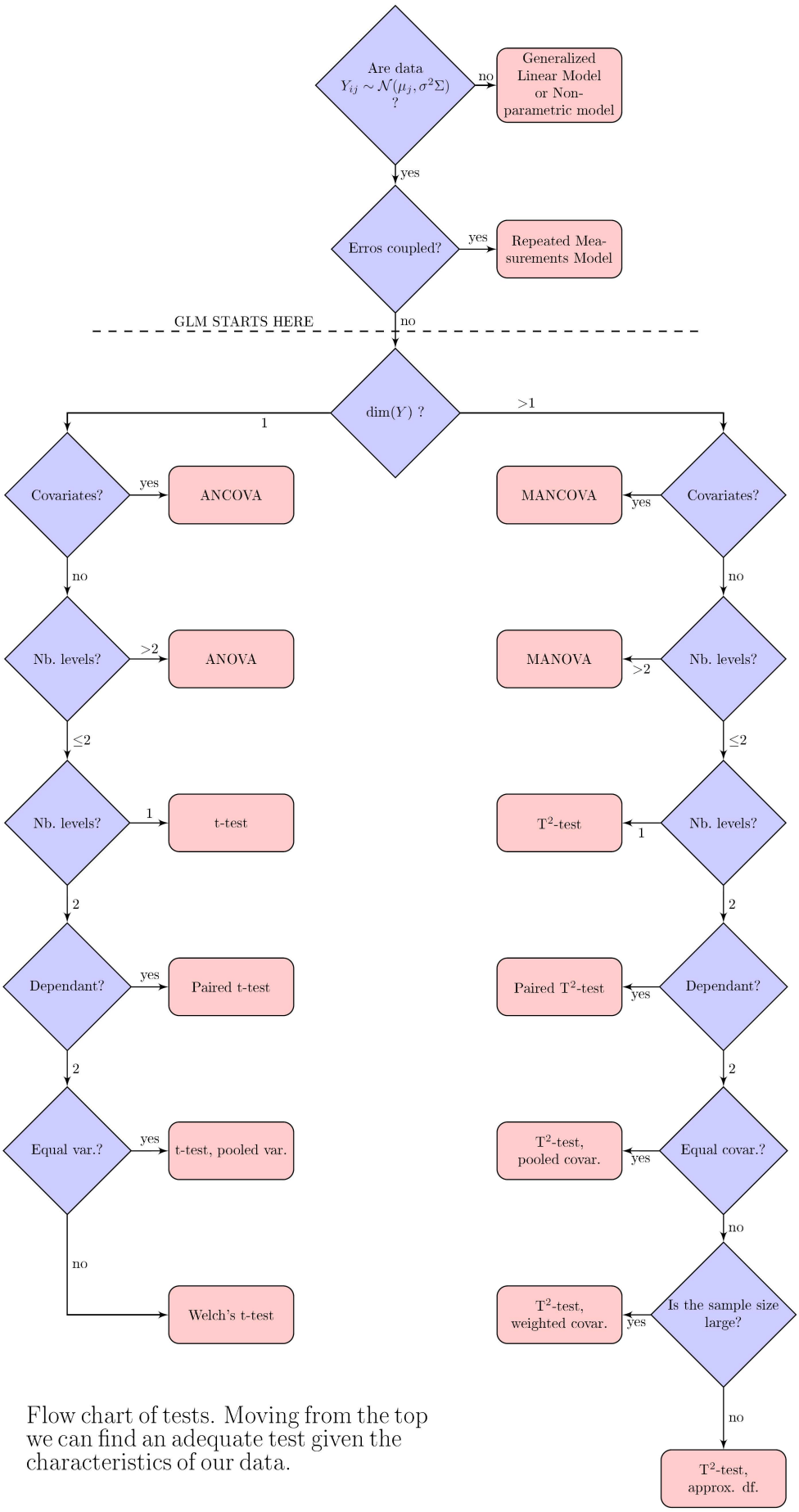
4.1 General Linear Model

The general linear model is a powerful framework that encompasses many different tests, see the flowchart on page 37. The basic formulation is

$$\mathbf{Y} = \mathbf{XB} + \boldsymbol{\epsilon} \tag{4.1}$$

where $\mathbf{Y}$ is a matrix of measurements, $\mathbf{X}$ is our model/design matrix, $\mathbf{B}$ is a matrix of coefficients to be estimated and $\boldsymbol{\epsilon}$ is a matrix of noise or errors. The errors are assumed to be normally distributed $\mathcal{N}(\mathbf{0}, \sigma^2\boldsymbol{\Sigma})$ with covariance $\boldsymbol{\Sigma}$. We will return to the model where $\boldsymbol{\Sigma}$ is structured with unknown parameters in the section about Repeated Measurement Models 4.2.

In the following only the one-way tests are shown. The review is based on [Johnson et al., 2007] and [Ersbøll and Conradsen, 2012].



Flow chart of tests. Moving from the top we can find an adequate test given the characteristics of our data.

Table 4.1: Analysis of Variance table

Source of variation	Sums of Squares	Degrees of freedom
Treatments	$\sum_{i=1}^{k} n_i(\bar{y}_i - \bar{y})^2$	$k-1$
Error	$\sum_{i=1}^{k}\sum_{i=j}^{n_i}(y_{ij} - \bar{y}_i)^2$	$\sum n_i - k$
Total	$\sum_{i=1}^{k}\sum_{i=j}^{n_i}(y_{ij} - \bar{y})^2$	$\sum n_i - 1$

4.1.1 Analysis of variance

Analysis of variance (ANOVA) is a subset of the GLM. Here the left hand side is univariate and we have i treatment effects:

$$Y_{ij} = \mu + \tau_i + \epsilon_{ij} \tag{4.2}$$

where Y_{ij} is the measurement from sample j belonging to treatment i for a total of N samples, μ is the overall mean, τ_i is the treatment effect of treatment i and ϵ_{ij} is the errors that are $\mathcal{N}(\mathbf{0}, \sigma^2)$ distributed. Let n_i denote the number of samples in treatment i, it then follows that $\sum_i n_i\tau_i = 0$.

The null-hypothesis, that we see no effects of the k treatments:

$$H_0 : \tau_1 = \tau_2 = \ldots = \tau_k \tag{4.3}$$

can then be tested be dividing the variation into three effects. We consider the sum of squares (SS) of the effect, the error and the total. Those are given by:

$$SS_{\text{treatment}} = \sum_{i=1}^{k} n_i(\bar{y}_i - \bar{y})^2 \tag{4.4}$$

$$SS_{\text{error}} = \sum_{i=1}^{k}\sum_{i=j}^{n_i}(y_{ij} - \bar{y}_i)^2 = SS_{\text{total}} - SS_{\text{effect}} \tag{4.5}$$

$$SS_{\text{total}} = \sum_{i=1}^{k}\sum_{i=j}^{n_i}(y_{ij} - \bar{y})^2 \tag{4.6}$$

It is customary to write this as a table, see Table 4.1.

From these we construct the F-test, which will reject the null-hypothesis at level α if fulfilled.

$$\frac{SS_{\text{treatment}}/(k-1)}{SS_{\text{error}}/(N-k)} > F(k-1, N-k)_{1-\alpha} \tag{4.7}$$

where F is the cumulative F-distribution $F(v_1, v_2)_{1-\alpha}$ with significance level $1-\alpha$ and v_1, v_2 degrees of freedom.

This methodology was applied in paper B, C, and D. In the first to assess various models describing the degradation of two types of gels. In the second to assess the significance of various factors that could influence the measured activity of ^{106}Ag used for in-vivo dosimetry. In the latter to assess the impact of physical exercise on the brain volume of alzheimers patients.

4.1.2 t-test

The t-test can be formulated as a subset of the ANOVA model and hence the GLM, but may be used completely independent here off. Here we are restricted to max two levels or treatments. While it can formulated as a GLM it is customary to use a more direct notation and tests. For the 2 sample case ($\mathbf{y}_1$,$\mathbf{y}_2$) with the same variance, we can test the hypothesis $H_0 : \mu_1 = \mu_2$ against $H_1 : \mu_1 \neq \mu_2$ with the test statistic:

$$\left| \frac{\bar{y}_1 - \bar{y}_2}{s_p \sqrt{1/n_1 + 1/n_2}} \right| > t(n_1 + n_2 - 2)_{1-\alpha/2} \tag{4.8}$$

where $\bar{y}$ is the mean and s_p is the pooled sample standard deviation given by:

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} \tag{4.9}$$

where s_1 and s_2 are the sample standard deviations, and n_1, n_2 the sample sizes. $t(v)_{1-\alpha/2}$ is the cumulative t-distribution with significance level α and degrees of freedom v.

This test was applied in paper D to assess whether significant differences were present between a test and a control group of patients.

4.1.3 Welch's t-test

In the 2 sample case ($\mathbf{y}_1$,$\mathbf{y}_2$) we again wish to test the hypothesis $H_0 : \mu_1 = \mu_2$ against $H_1 : \mu_1 \neq \mu_2$. If we cannot assume equal variances of the samples we use Welch's t-test instead:

$$\left| \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \right| > t(DF)_{1-\alpha/2} \tag{4.10}$$

where the degrees of freedom DF is estimated by the Welch-Satterthwaite equation:

$$DF = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{(s_1^2/n_1)^2/(n_1-1) + (s_2^2/n_2)^2/(n_2-1)} \tag{4.11}$$

This test was applied in paper D to assess whether significant differences were present between a test and a control group of patients. The test was applied, when the data could not be assumed to have equal variance.

4.1.4 Multivariate analysis of variance

Multivariate analysis of variance (MANOVA) can be considered as a subset of the GLM, but may be used independently. We look at the one-way test, but it can be easily extended to two or more effects. Here we have multi dimensional measurements on the left hand side:

$$\mathbf{Y}_{ij} = \boldsymbol{\mu} + \boldsymbol{\tau}_i + \boldsymbol{\epsilon}_{ij} \tag{4.12}$$

where $\mathbf{Y}_{ij}$ is the measurements from sample j belonging to treatment i for a total of N samples, $\boldsymbol{\mu}$ is the overall mean, $\boldsymbol{\tau}_i$ is the treatment effect of treatment i and $\boldsymbol{\epsilon}_{ij}$ is a matrix of errors that is $\mathcal{N}(\mathbf{0}, \sigma^2\boldsymbol{\Sigma})$ distributed. It follows that $\sum_i \boldsymbol{n}_i\boldsymbol{\tau}_i = \mathbf{0}$.

The SS used in the ANOVA generalises to the multivariate case. We look at the between ($\mathbf{B}$), within ($\mathbf{W}$), and total ($\mathbf{T}$) variation in the data. That is the variation between the treatments (eqv. to $SS_{\text{treatment}}$), the variation within the treatment groups (eqv. to SS_{error}) and the total variation in the data (eqv. to SS_{total}). Given k treatments we get:

$$\mathbf{B} = \sum_{i=1}^{k} n_i(\bar{\mathbf{Y}}_i - \bar{\mathbf{Y}})(\bar{\mathbf{Y}}_i - \bar{\mathbf{Y}})^{\mathrm{T}} \tag{4.13}$$

$$\mathbf{W} = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(\mathbf{Y}_{ij} - \bar{\mathbf{Y}}_i)(\mathbf{Y}_{ij} - \bar{\mathbf{Y}}_i)^{\mathrm{T}} \tag{4.14}$$

$$\mathbf{T} = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(\mathbf{Y}_{ij} - \bar{\mathbf{Y}})(\mathbf{Y}_{ij} - \bar{\mathbf{Y}})^{\mathrm{T}} \tag{4.15}$$

Instead of the F-test we can use Wilks Λ (Andersons U):

$$\Lambda = \mathrm{U} = \frac{|\mathbf{W}|}{|\mathbf{T}|} = \frac{|\mathbf{W}|}{|\mathbf{B}+\mathbf{W}|} \leq \Lambda(p, k-1, N-k)_{1-\alpha} \tag{4.16}$$

where $|\ \ |$ denotes determinant, and $\Lambda(p, k-1, N-k)_{1-\alpha}$ is the cumulative Λ-distribution at significance level α, and p is the dimensionality of $\mathbf{Y}$.

This methodology was applied in paper D, to assess the impact of physical exercise in brain regions of alzheimers patients.

4.1.5 Hotellings T^2

Equivalent to the situation with ANOVA, the test is equivalent to a simpler formulation when we only have 2 levels. For this situation we use Hotelling T^2-test. Given two samples $(\mathbf{Y}_{1i}, i = 1, \cdots, n_1)$, $(\mathbf{Y}_{2j}, j = 1, \cdots, n_1)$ and assuming equal covariance the formulation is:

$$T^2 = \left(\frac{1}{n_1} + \frac{1}{n_2}\right)^{-1/2} (\bar{\mathbf{Y}}_\mathbf{1} - \bar{\mathbf{Y}}_\mathbf{2})^{\mathrm{T}}\, \mathbf{S}_{\text{pooled}}^{-1} \left(\frac{1}{n_1} + \frac{1}{n_2}\right)^{-1/2} (\bar{\mathbf{Y}}_\mathbf{1} - \bar{\mathbf{Y}}_\mathbf{2}) \quad (4.17)$$

where n_1, n_2 are the sample sizes and $\mathbf{S}_{\text{pooled}}$ is the pooled covariance given by:

$$\mathbf{S}_{\text{pooled}} = \frac{n_1 - 1}{n_1 + n_2 - 2}\mathbf{S}_\mathbf{1} + \frac{n_2 - 1}{n_1 + n_2 - 2}\mathbf{S}_\mathbf{2} \quad (4.18)$$

where $\mathbf{S}_\mathbf{1}$ and $\mathbf{S}_\mathbf{1}$ are the sample covariances. We reject the null-hypothesis, that $H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$ against $H_1 : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2$ if:

$$T^2 > c^2 = \frac{(n_1 + n_2 - 2)p}{(n_1 + n_2 - p - 1)} F(p, n_1 + n_2 - 1)_\alpha \quad (4.19)$$

where F is the cumulative F-distribution at significance level α with p, n_1+n_2-1 degrees of freedom.

Like the t-test the T^2-test can be extended to unequal covariances for the 2 samples, see e.g. [Johnson et al., 2007].

Indirectly this test was applied in paper D, to assess the impact of physical exercise in brain regions of alzheimers patients. Although the test were formulated as a MANOVA 4.1.4, this reduces to Hotellings T^2 when we only have two classes.

4.1.6 Estimation in the General Linear Model

To estimate the parameters in the GLM, we rely on either the inverse or the pseudo inverse of the model matrix. We recall that the GLM was:

$$\mathbf{Y} = \mathbf{XB} + \boldsymbol{\epsilon} \tag{4.20}$$

where $\mathbf{Y}$ is a matrix of measurements, $\mathbf{X}$ is our model matrix, $\mathbf{B}$ is a matrix of coefficients to be estimated and $\boldsymbol{\epsilon}$ is a matrix of noise or errors.

Without loss of generality, we look at the univariate case, i.e. Y is a vector distributed $\mathrm{Y} \sim \mathcal{N}(\mathbf{XB}, \sigma^2)$ and B is a vector of coefficients to be estimated. B can be estimated using a least squares approach:

$$\hat{\mathrm{B}} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathrm{Y} \tag{4.21}$$

If $\mathbf{X}^{\mathrm{T}}\mathbf{X}$ is not full rank we use the pseudo inverse instead.

In practice SAS, R or another statistical language is used for the computations. SAS has been used in this thesis and in PROC GLM it utilises the generalised g2-inverse of $\mathbf{X}^{\mathrm{T}}\mathbf{X}$ [SAS, 2008].

As long as the experiment is balanced, no problems arises. However, as soon as we have missing data or for some other reason have an unbalanced experiment our coefficients and thus sums of squares are not uniquely identifiable. In that context different types of SS are reported. The following review is based on [Iacobucci, 1995, Driscoll and Borror, 2000], but the SAS notation is used [SAS, 2008]. We distinguish between four types of layout, with the notation $n(\ ,\ )$ denoting number of samples in a cell:

1 Balanced layout: Equal cell frequencies, $n(i,j)$ constant for all i,j combinations
2 Proportional layout: $\frac{n(i,j)}{n(i,l)} = \frac{n(k,j)}{n(k,l)}$ for all i,j,k,l combinations
3 Unbalanced layout: $n(i,j) > 0$ for all i,j, but [2] does not hold for all i,j,k,l combinations
4 Missing cells: $n(i,j) = 0$ for some i,j

We will use a model with an overall mean μ, effects A, B, and interaction AB:

$$Y = \mu + A + B + AB + \epsilon \tag{4.22}$$

to describe the different types of SS. The notation will be as follows: $SS(A|B) = SS(B,A) - SS(B)$.

4.1.6.1 Type I

Type I is a sequential estimation, and thus dependant on the order of the effects in the model.

$$SS(A|\mu)$$
$$SS(B|\mu, A)$$
$$SS(AB|\mu, A, B)$$

A nice property of Type I is that the effects SS sum up to the model SS.

Appropriate for:

- purely nested models (specified in the proper sequence)
- polynomial regression models (in the proper sequence)

4.1.6.2 Type II

Type II estimates the SS after all other effects have been added to the model except those that contain the given effect.

$$SS(A|\mu, B)$$
$$SS(B|\mu, A)$$
$$SS(AB|\mu, A, B)$$

Appropriate for:

- any main-effects model
- any pure regression model
- an effect not contained in any other effect (regardless of the model)

4.1.6.3 Type III

Type III estimates the SS after controlling for all other effects.

$$SS(A|\mu, B, AB)$$
$$SS(B|\mu, A, AB)$$
$$SS(AB|\mu, A, B)$$

Appropriate for:

- Unbalanced layout
- Missing cells

4.1.6.4 Type IV

Similar to Type III but handles missing cells differently.

Appropriate for:

- Unbalanced layout
- Missing cells

To summarise the different properties see table 4.2.

Table 4.2: Equalities in Type I-IV SS

Effect	Balanced	Proportional	Unbalanced	Missing cells
A	I=II=III=IV	I=II, III=IV	III=IV	(all unequal)
B	I=II=III=IV	I=II, III=IV	I=II, III=IV	I=II
AB	I=II=III=IV	I=II=III=IV	I=II=III=IV	I=II=III=IV

The literature favours using Type III in general. However, in practical work one tries to obtain the simplest possible model that describe the observations. Then it will often be necessary to perform a sequential testing based on successively simpler and simpler models. In that case one must set up the test statistics "manually".

This section is included for completeness and Type III was used in the included works, e.g. in paper C to assess the significance of various factors that could influence the measured activity of ^{106}Ag used for in-vivo dosimetry.

4.2 Repeated Measurements Model

In the repeated measurements models (RMM) we use the same structure as the in the GLM:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{4.23}$$

where $\mathbf{y}$ is a vector of measurements, $\mathbf{X}$ is our model matrix, $\boldsymbol{\beta}$ is a vector of coefficients to be estimated and $\boldsymbol{\epsilon}$ is a vector of noise or errors. The change

is, that the errors now are assumed to have a specific structure. They are still normally distributed $\mathcal{N}(\mathbf{0}, \sigma^2\mathbf{\Sigma})$ but the covariance $\mathbf{\Sigma}$ now has a specific structure and additional parameters that has to be estimated.

By letting the errors be correlated, we can use the information in a longitudinal study. A subject with a higher value than the mean, is likely to also have a higher value on the next measurement as the measurements are on the same subject and thus correlated.

PROC GLM allows for RMM to some degree, but a more flexible framework is found in PROC MIXED [SAS, 1999]. The documentation for PROC MIXED is used in the following. As opposed to PROC GLM the procedure PROC MIXED uses a restricted maximum likelihood estimate (REML) instead of a generalised inverse.

As PROC MIXED also allows for random effects the model changes to:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\epsilon} \tag{4.24}$$

where $\mathbf{Z}$ is a known matrix and $\boldsymbol{\gamma}$ is a vector of unknown random effects to be estimated.

Further we have the assumptions that $\boldsymbol{\gamma}$ and $\boldsymbol{\epsilon}$ are normally distributed with:

$$\mathrm{E}\begin{bmatrix}\boldsymbol{\gamma}\\ \boldsymbol{\epsilon}\end{bmatrix} = \begin{bmatrix}\mathbf{0}\\ \mathbf{0}\end{bmatrix} \tag{4.25}$$

$$\mathrm{Var}\begin{bmatrix}\boldsymbol{\gamma}\\ \boldsymbol{\epsilon}\end{bmatrix} = \begin{bmatrix}\mathbf{G} & \mathbf{0}\\ \mathbf{0} & \mathbf{R}\end{bmatrix} \tag{4.26}$$

The general linear model is then simply a special case with $\mathbf{Z} = (\mathbf{0})$ and $\mathbf{R} = \sigma^2\mathbf{\Sigma}$.

The variance of $\mathbf{y}$ is then $\mathbf{V} = \mathbf{Z}\mathbf{G}\mathbf{Z}^{\mathrm{T}} + \mathbf{R}$. The restricted maximum likelihood can then be formulated as:

$$\begin{aligned}\ell(\mathbf{G}, \mathbf{R}) = &-\frac{1}{2}\log|\mathbf{V}| - -\frac{1}{2}\log|\mathbf{X}^{\mathrm{T}}\mathbf{V}^{-1}\mathbf{X}|\\ &-\frac{1}{2}\mathbf{r}^{\mathrm{T}}\mathbf{V}^{-1}\mathbf{r} - \frac{n-p}{2}\log(2\pi)\end{aligned} \tag{4.27}$$

where $\mathbf{r} = \mathbf{y} - \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{V}^{-1}\mathbf{X})^{-}\ \mathbf{X}^{\mathrm{T}}\mathbf{V}^{-1}\mathbf{y}$, p is the rank of $\mathbf{X}$, and $(^{-})$ denotes the generalised inverse. A rather ugly equation but nonetheless solvable using a ridge stabilised Newton-Raphson algorithm. When we have estimated $\hat{\mathbf{R}}$, $\hat{\mathbf{G}}$ and thus $\hat{\mathbf{V}}$, the parameters $\mathbf{B}$ and $\boldsymbol{\gamma}$ can simply be found using least square:

$$\hat{\mathbf{B}} = (\mathbf{X}^{\mathrm{T}}\hat{\mathbf{V}}^{-1}\mathbf{X})^{-}\mathbf{X}^{\mathrm{T}}\hat{\mathbf{V}}^{-1}\mathbf{Y} \tag{4.28}$$

$$\hat{\gamma} = \hat{\mathbf{G}}\mathbf{Z}^{\mathrm{T}}\hat{\mathbf{V}}^{-1}(\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}}) \tag{4.29}$$

The special case for $\hat{\mathbf{G}}$ singular is treated in [SAS, 1999].

The model was used in paper B to assess degradation over time in gels. The gels were placed in mice that were measured repeatedly.

4.2.1 Covariance structure in the errors

A wide variety of covariance structures are possible in PROC MIXED. In this thesis only the autoregressive structure was used - see paper B where a RMM was used to estimate degradation of two types of gels. Different subjects are assumed uncorrelated, while the measurements within one subject has a time dependent correlation. This gives the following variance-covariance matrix:

$$\mathbf{D}(\epsilon_{jk}) = \sigma^2 \begin{bmatrix} 1 & \cdots & \rho^{|t(k)_{1j} - t(k)_{Ij}|} \\ \vdots & \ddots & \vdots \\ \rho^{|t(k)_{1j} - t(k)_{Ij}|} & \cdots & 1 \end{bmatrix} \tag{4.30}$$

where t is time and I is the last observation for subject j group k. It has the interpretation that the correlation between two measurements decrease with time between the measurements. This seems like a sensible assumption.

4.3 Assumption tests

All of the tests presented in the previous sections hinge on the assumption of normality. Several of them also depend on the variance or covariance to be equal. Finally, a given model can be evaluated be inspecting the residuals.

In this section we will start with the influence statistics and residuals, before moving to more rigorous tests of equality of variance or covariance.

For this section it is convenient to define the hat matrix:

$$\mathbf{H} = \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-}\mathbf{X}^{\mathrm{T}} \tag{4.31}$$

where $\mathbf{X}$ is our model matrix.

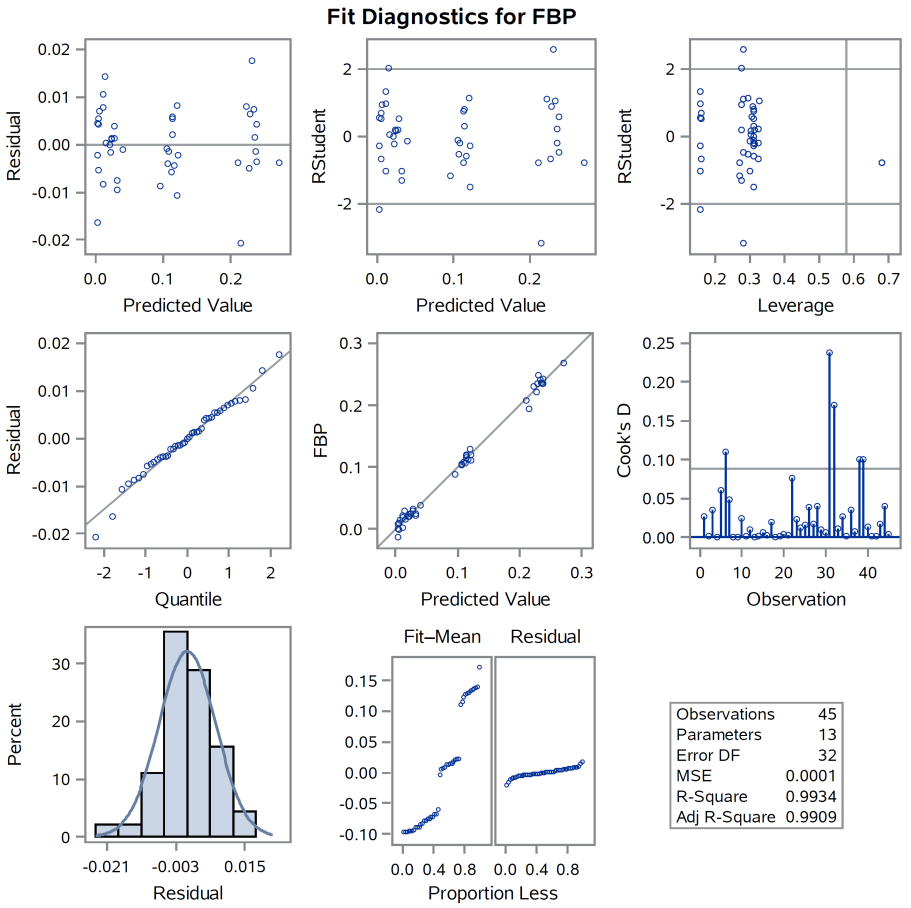


Figure 4.1: Example of fit diagnostics from SAS

4.3.1 Influence statistics and residuals

An example of the standard fit diagnostics output by SAS can be seen in fig. 4.1.

These were used to assess model adequacy in papers B, C, and D. For examples see the sections below and appendix C to paper B.

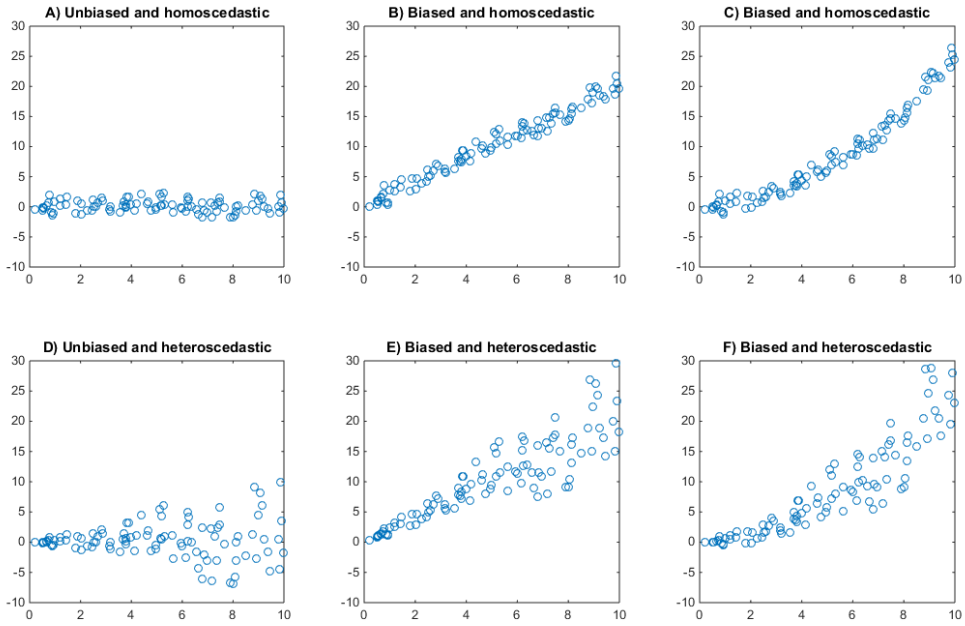


Figure 4.2: Example of residual plots in different situations

4.3.1.1 Residual plots

The upper left figure in fig. 4.1 shows a residual plot. This can be used to diagnose any potential lacks in the fitted model. Six examples are given in fig. 4.2. The residuals in A) are quite like those in fig. 4.1. We see no trend in the residuals, and it appears we have an unbiased estimate. In subplot B) we need a linear and in C) a quadratic term in our model, to fit the trend adequately. In subplots D)-F) we have a growing variance. In those situations we would often resort to a variance stabilising transformation of the data.

4.3.1.2 RSTUDENT

RSTUDENT is a normalisation of the residual and is defined as:

$$\text{RSTUDENT}_i = \frac{r_i}{\hat{\sigma}(i)\sqrt{1-h_{ii}}} \tag{4.32}$$

where $\hat{\sigma}(i)^2$ is the variance estimated without the ith observation and h_{ii} is element i, i in the hat matrix. If RSTUDENT is larger than $\pm$ 2–3 the observation warrants extra scrutiny.

Examples can be seen in fig. 4.1 in the top middle and right figures.

4.3.1.3 Leverage

Leverage is a measure of how far from the mean the ith independent variable is and is defined as:

$$h_{ii} = \mathbf{X}_i(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-}\mathbf{X}_i^{\mathrm{T}} \tag{4.33}$$

Since such a measure or observation - given a high leverage - can force or influence the regression to lie closer to this observation they are worth considering. As a rule of thumb observations with a leverage $> 2p/n$ should be investigated, where p is the parameters in the model and n the total number of observations.

An example can be seen in fig. 4.1 in the top right figure.

4.3.1.4 Cook's D

Cook's distance is a measure of the effect of deleting a given observation and is defined as:

$$D_i = \frac{1}{p\hat{\sigma}^2}(\hat{\mathbf{B}} - \hat{\mathbf{B}}(i))^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{X}(\hat{\mathbf{B}} - \hat{\mathbf{B}}(i)) \tag{4.34}$$

where $\hat{\mathbf{B}}(i)$ is the parameter estimates without observation i.

An example can be seen in fig. 4.1 in the middle right figure.

4.3.1.5 Fit-mean and Residual plot

In bottom middle in fig. 4.1 a Fit-Mean (pronounced fit minus mean) and residual plot can be seen.

It is used to compare the spread of the fit to the spread of the residuals. If the spread of the residuals is not much smaller than the spread of the mean, our model is not very good and much of the variance is not explained by the model.

4.3.2 Normality tests

While more rigorous normality tests exist it is often sufficient to inspect residual plots. Examples of these tests are [Shapiro and Wilk, 1965, Anderson, 1962, Anderson and Darling, 1952, Lilliefors, 1967], we will, however, limit ourselves to the residual plots.

The methods below were used to assess normality in paper B, C, and D.

4.3.2.1 Q-Q plot

A Q-Q plot can be seen in the middle left of fig. 4.1. It is generated by plotting the ordered residuals against the quantiles of a normal distribution. If the residuals are normal they will follow a straight line. In the example in fig. 4.1, the residuals look normal except perhaps for the two lowest observations.

4.3.2.2 Histogram plot

A Histogram plot with normal distribution overlaid can be seen in the bottom left of fig. 4.1. The distribution is clearly unimodal, and we see nothing that would lead us to reject the normality assumption.

4.3.3 Variance tests

When considering the most common case for a general linear model, namely $\Sigma = I$, equality of variance or homoscedascity becomes an assumption. As mentioned earlier, we can evaluate homoscedascity from the residual plots but more rigorous tests are available when we are in doubt.

4.3.3.1 Univariate: Levene

Levene's test for equal variance is given by [Brown and Forsythe, 1974]:

$$W = \frac{\sum_i n_i(\bar{z}_i - \bar{z}_{\cdot\cdot})^2/(g-1)}{\sum_i \sum_j (\bar{z}_i - \bar{z}_{i\cdot})^2 / \sum_i (n_i - 1)} \tag{4.35}$$

where:

$$z_{ij} = |y_{ij} - \bar{y}_i| \tag{4.36}$$

$$\bar{z}_i = \sum_j z_{ij}/n_i \tag{4.37}$$

$$\bar{z}_{..} = \sum_i \sum_j z_{ij} / \sum_i n_i \tag{4.38}$$

The critical value is found by using the cumulative F-distibution: $F(g-1, \sum_i (n_i - 1))_{1-\alpha}$. Unfortunately the test is very sensitive to departures from normality.

This test was used in paper C to assess variance between the different dose levels delivered to the Ag-nanogel. See paper C supplementary material.

4.3.3.2 Multivariate: Box's M

Box's M test [Box, 2015, Johnson et al., 2007] is used to test for equal variance in the multivariate case. It is available in SAS through the PROC DISCRIM procedure. It is also sensitive to departures from normality. In addition it is quite powerful and may detect differences that are trivial.

This was used to assess possible differences in variance between control and test-group in the alzheimers study in paper D.

4.4 Model Selection

In the GLM framework we can test sequentially for nested submodels by using the F-test. However, it is often the case that models are not nested or we do not test sequentially. In these situations where no direct test is available other criteria are more fitting.

The information criteria described below was used in paper B to select between various RMMs describing the degradation of two types of gel.

4.4.1 R^2 and adjusted R^2

A very useful - although a bit crude - measure for a model, is the multiple correlation coefficient [Ersbøll and Conradsen, 2012]:

$$R^2 = \frac{SS_{\text{total}} - SS_{\text{error}}}{SS_{\text{total}}} = \frac{SS_{\text{effect}}}{SS_{\text{Total}}} \tag{4.39}$$

As can be seen R^2 is a measure for the amount of variance explained by our model.

This - in itself - is a poor measure of how good a model is. R^2 tends to increase as parameters are added to the model, and a very complex model can achieve a high R^2 - but we need to weigh that against the complexity. For that we have the adjusted R^2:

$$R^2_{adj} = 1 - \frac{(n-i)(1-R^2)}{n-p} \tag{4.40}$$

where n is the number of observations, p is the number of parameters, and i is equal to 1 if an intercept is included in the model and 0 otherwise.

4.4.2 Information Criteria

The information criteria are a relative measure of the quality of a model. It provides a trade-off between the information lost in a given model and the complexity of said model

4.4.2.1 Akaikes information criterion

Akaikes information criterion [Akaike, 1974] is defined as:

$$\text{AIC} = 2k - 2\log(L) \tag{4.41}$$

where k is the number of parameters and L is the maximised values of the likelihood function of the model. The preferred model is then the one that has the lowest AIC. As we have finite sample sizes a correction is needed, the AICC [Hurvich and Tsai, 1991]:

$$\text{AICC} = \text{AIC} + \frac{2k(k+1)}{n-k-1} \tag{4.42}$$

where n is the sample size and k is again the number of parameters.

If we have equal variance among the models, AIC reduces to a very convenient form:

$$\text{AIC} = 2k + n\ \log(\text{RSS}/n) + C_1 = 2k + n\ \log(\text{RSS}) + C_2 \tag{4.43}$$

where RSS is the residual sum of squares and C_1 and C_2 are constants across the models and can be ignored.

As AICC converges to AIC as the sample size increases, it should generally be preferred [Burnham, 2004].

4.4.2.2 Bayesian information criterion

The Bayesian information criterion was originally formulated by [Schwarz, 1978] but is now used in a slightly different formulation [Burnham, 2004] is defined as:

$$\text{BIC} = -2\ \log(L) + k\ \log(n) \tag{4.44}$$

where k is the number of parameters, L is the maximised values of the likelihood function of the model and n the sample size. BIC penalises models more heavily than AIC but is comparable to AICC.

Like AIC it reduces to a simple form when the variance is equal over the models:

$$\text{BIC} = n\ \log(\text{RSS}/n) + k\ \log(n) \tag{4.45}$$

Whether AICC or BIC should be preferred is still not settled.

4.5 Non-parametric methods

When presented with non-normal data we may be able to model and test in the Generalised Linear Model. However, even that is not always possible. To accommodate the general case, where we make no - or only weak - assumptions on the distribution of data we have the non-parametric models.

4.5.1 Mann-Whitney-U

For univariate distribution the Mann-Whitney-U test [Johnson and Gibbons, 1973], equivalent to the Wilcoxon rank-sum test, can be used to test whether two samples come from the same population.

Given two samples $\mathbf{X} = X_1, X_2, \ldots, X_m$ and $\mathbf{Y} = Y_1, Y_2, \ldots, Y_n$ we start by ordering the samples in ascending order and then merge them into a single sequence of $m + n = N$ variables. We can then form an indicator variable:

$$D_{ij} = \begin{cases} 1 & \text{if } Y_j < X_i \\ & \quad \text{for } i = 1 \ldots m,\ j = 1 \ldots n \\ 0 & \text{if } Y_j > X_i \end{cases} \tag{4.46}$$

and the test statistic is then:

$$U = \sum_{i=1}^{m} \sum_{j=1}^{n} (1 - D_{ij}) \tag{4.47}$$

The critical value for U can be tabulated for small values of m and n or approximated by a normal distribution for larger values.

This test was applied in paper A to test for possible differences in variance - and thus inhomogeneity - between two different formulations of a Ag-nanogel. It was also used to assess whether significant differences were present between a test and a control group of patients in paper D.

4.5.2 Multivariate rank tests

For the multivariate case the Oja rank tests can be used [Oja and Randles, 2004, Oja, 2010]. Naturally, they are somewhat more complicated than the univariate case but quite elegant.

Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ be identically, independently distributed matrices of p dimensional measurements sampled from $F(\mathbf{x} + \boldsymbol{\mu})$ with a total of N samples. $F(\cdot)$ represents a continuous p-dimensional distribution with location parameters $\boldsymbol{\mu}$.

We define the spatial sign function:

$$Sign(x) = \begin{cases} \|\mathbf{x}\|^{-1}\mathbf{x} & \mathbf{x} \neq \mathbf{0} \\ 0 & \mathbf{x} = \mathbf{0} \end{cases} \tag{4.48}$$

The multivariate rank is then:

$$\mathbf{S}_{ij} = Sign(A_x(\mathbf{X}_i - \mathbf{X}_j)), \qquad i, j = 1, \ldots, n \tag{4.49}$$

where A_x is a transformation matrix, that we will return to shortly. From this we form the centered rank:

$$\mathbf{R}_i = \frac{1}{n} \sum_j \mathbf{S}_{ij} \tag{4.50}$$

Returning to A_x we choose it to make the rank affine invariant. This can be achieved by satisfying the property:

$$p\frac{1}{n}\sum_i \mathbf{R}_i\mathbf{R}_i^{\mathrm{T}} = (\frac{1}{n}\sum_i \mathbf{R}_i^{\mathrm{T}}\mathbf{R}_i)I \tag{4.51}$$

where I is the identity matrix.

The transformation can be found iteratively. Let $A_x^{\mathrm{T}}A_x = V_x^{-1}$, we then initialise $V = I$ and use the update:

$$V_{+1} = \frac{p}{\frac{1}{n}\sum_i \mathbf{R}_i^{\mathrm{T}}\mathbf{R}_i}V^{1/2}\frac{1}{n}\sum_i \mathbf{R}_i\mathbf{R}_i^{\mathrm{T}}V^{1/2} \tag{4.52}$$

Unfortunately the convergence of this algorithm has yet to be proofed, but it seems to work well in practice.

The several sample spatial rank test with c groups is then:

$$U^2 = \frac{p}{\frac{1}{n}\sum_i \mathbf{R}_i^{\mathrm{T}}\mathbf{R}_i}\sum_{i=1}^{c} n_i\|\bar{\mathbf{R}}_i\|^2 \tag{4.53}$$

The p-value can be found by permuting the data:

$$E_\gamma\left[U_\gamma^2 \geq U^2\right] \tag{4.54}$$

where $\gamma = (\gamma_1, \ldots, \gamma_N)$ is a uniform distributed sample of the $N!$ possible permutations and U_γ^2 is the test statistic of the permuted sample.

This test was used in paper D to confirm the results obtained from a Hotellings T^2-test, as there were several outliers present in the data.

4.6 Variance stabilising transforms

If - as we saw in fig. 4.2 (d-f) - the data are heteroscedastic, we can try to stabilise the variance by using various transforms. Some of the most commonly employed are to take the logarithm for log-normally data and the square root for Poisson distribution. A more general transformation is the Box-Cox transformation.

4.6.1 Box-Cox

The Box-Cox transformation [Box and Cox, 1964] in addition to being variance stabilising, is also able to correct for skewness and make the data more normal.

The transformation is defined as:

$$y'_\lambda = \begin{cases} \frac{y^\lambda - 1}{\lambda} & \lambda \neq 0 \\ \log(y) & \lambda = 0 \end{cases} \tag{4.55}$$

the question is now how to choose the parameter λ. For a single variable we do that by maximising the log-likelihood function [Johnson et al., 2007]:

$$L = -\frac{n}{2}\log(\sigma^2(\frac{y^\lambda - 1}{\lambda})) - (\lambda - 1)\sum_y(\log(y)) \tag{4.56}$$

In the GLM framework, the log-likelihood function can be extended to incorporate the design matrix $\mathbf{X}$.

The Box-Cox transform was used in paper A for variance stabilisation allowing for comparison of variance betweens gels with different contrast level.

4.7 Correlation tests

For test of structure in correlation and covariance matrices PROC CALIS [Sas, 2007] was used. It allows for the specification of a constrained covariance matrix to be tested against the unconstrained matrix.

Let p be the number of variables, $\mathbf{S}$ the sample $p \times p$ covariance matrix, $\bar{x}$ the vector of sample means, $\mathbf{\Sigma}$ the fitted constrained covariance matrix and μ the vector of fitted means.

We now minimise a discrepancy function, which when using the likelihood principle is:

$$F = \mathrm{Tr}(\mathbf{S}\mathbf{\Sigma}^{-1}) - p + \log(|\mathbf{\Sigma}|) - \log(|\mathbf{S}|) \tag{4.57}$$

The test statistic is then simply:

$$C = (N - k - l) \cdot F_{min} \tag{4.58}$$

where N is the number of observations, k is the number of groups and l is the number of variables used as covariates that we have partialled against, i.e variables we have controlled for [1]. This value can then be tested in cumulative $\chi^2(C, m)$ with degrees of freedom m which is the number of parameters constrained in $\mathbf{\Sigma}$.

[1] The SAS documentation does not mention that we need to subtract the number of variables partialled against, which can be rather confusing...

The covariance test was used in paper D to assess correlation between several brain regions and a single test score. The test score could e.g. be cognitive performance and it was of interest whether this was couples to the development in certain regions of the brain.

4.8 Performance measures

To assess performance of a segmentation or classification, we can use the confusion/error matrix [Fawcett, 2006, Labatut and Cherifi, 2011, Liu et al., 2007]:

Table 4.3: Confusion Matrix

		Actual	
		Positive	Negative
Predicted	Positive	True Positive (TP)	False Positive (FP)
	Negative	False negative (FN)	True negative (TN)

Based on these four labels we can generate a variety of measures:

Table 4.4: Performance measures

Performance measures	
sensitivity or true positive rate (TPR)	$\frac{TP}{TP+FN}$
specificity (SPC) or true negative rate (TNR)	$\frac{TN}{FP+TN}$
precision or positive predictive value (PPV)	$\frac{TP}{TP+FP}$
negative predictive value (NPV)	$\frac{TN}{TN+FN}$
fall-out or false positive rate (FPR)	$\frac{FP}{FP+TN}$
false discovery rate (FDR)	$\frac{FP}{FP+TP}$
miss rate or false negative rate (FNR)	$\frac{FN}{FN+TP}$
accuracy (ACC)	$\frac{TP+TN}{TP+FP+FN+TN}$
F1-score or Sørensen/Dice-score	$\frac{2TP}{2TP+FP+FN}$

Accuracy gives a good overall measure of the performance, if the categories are balanced, i.e. have approximately the same number of samples. If this not the case, the accuracy is a poor measure as we can obtain a high score by simply

always guessing at the dominant class. In cases like that one needs to consider other measures such as sensitivity, specificity and precision. A nice alternative to assessing performance by evaluating several measures is to use Matthew's correlation coefficient [Matthews, 1975]:

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{4.59}$$

which is a balanced measure that can be used even if the classes differ markedly in size.

This was used in paper E to quantify segmentation errors.

4.9 Classification

In the preceding chapter we went through various measures of how good performance a classifier has. Now for the actual classifiers. A classifier is a function that maps an input into discrete classes. An example is a simple threshold: If higher than: then one label. If not higher: then another label. We briefly go through some of the most prominent classification algorithms.

In this thesis only quite simple methods have found use. It is possible, but not probable, that the more advanced methods also presented here could have increased the performance.

Naive Bayes sec. 4.9.1 was used in paper F to classify between healthy, emphysematous and fibrotic lung tissue in mice. The quadratic classifier described in sec. 4.9.2 was then used to assess whether the classification could be improved by considering the correlation between the various features used. The rest are included for completeness.

4.9.1 Naive Bayes

Given a sample $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ we want to find $p(C_k|\mathbf{x})$ for each of the k classes. For that purpose we use Bayes' theorem [Pitman, 1993, Ersbøll and Conradsen, 2012]:

$$p(C_k|\mathbf{x}) = \frac{\pi_k p(\mathbf{x}|C_k)}{p(\mathbf{x})} \tag{4.60}$$

As the term $p(\mathbf{x})$ is equal over all classes k we can ignore it. π_k is the prior probability of a class. Often this will simply be $\frac{1}{k}$ if we have no additional information. $p(\mathbf{x}|C_k)$ can be inferred from our training data.

The name 'Naive' Bayes, derive from the rather naive assumption that the distribution of the features are independent.

4.9.2 Discriminant Analysis

Instead of looking at the individual frequencies as in Naive Bayes, we can make use of the correlation between the variables through Linear and Quadratic Discriminant Analysis [Hastie et al., 2009]. For each k out of a total of p classes, we model the distribution as:

$$f_k(x) = \frac{1}{(2\pi)^{p/2}|\Sigma_k|^{1/2}} e^{-\frac{1}{2}(x-\mu_k)\Sigma_k^{-1}(x-\mu_k)} \tag{4.61}$$

We can then apply Bayes theorem to get $P(C_k|x)$. For the case where Σ is equal over the the p classes we get the linear discriminant function:

$$\delta_k(x) = x^{\mathrm{T}}\Sigma^{-1}\mu_k - \frac{1}{2}\mu_k^{\mathrm{T}}\Sigma^{-1}\mu_k + \log(\pi_k) \tag{4.62}$$

If, however, the Σ is not equal over the classes we get the quadratic discriminant function:

$$\delta_k(x) = -\frac{1}{2}\log|\Sigma_k| - \frac{1}{2}(x-\mu_k)^{\mathrm{T}}\Sigma_k(x-\mu_k) + \log(\pi_k) \tag{4.63}$$

The sample x then belongs to the class with the highest value of δ. The Σ_k and μ_k are estimated from training data and π_k is the prior probability for class k.

The names linear and quadratic derive from the fact, that when solving and eq. 4.62 and eq. 4.63 respectively for x to get the decision boundary between each pair of classes we get a linear and a quadratic equation.

4.9.3 Logistic Regression

Logistic regression classifies by using the log-ratios between the probability of groups given the data. For two groups g_1 and g_2 we have:

$$\log\frac{P(G=g_1|X=x)}{P(G=g_2|X=x)} = \beta_0 + x\beta_x = 0 \tag{4.64}$$

Our decision boundary is the value where the probability of the group given the data is equal.

To find it we maximise the likelihood function of β:

$$\ell(\beta) = \sum_{i=1}^{N} \log P_{g_i}(x_i|\beta) \tag{4.65}$$

Logistic regression can easily be extended to more than two groups. It has fewer assumptions than LDA or QDA and is generally more robust. However, it breaks down if the classes are perfectly separable by a hyperplane. See [Hastie et al., 2009] for more details.

4.9.4 Random forest

Random forest is a collection of decision tress, which - together - forms a forest. The method aims at averaging many approximately unbiased but noisy trees to obtain low variance results. The following is based on [Hastie et al., 2009].

The procedure is as follows:

1. Draw a bootstrap sample from the training data
2. Grow a tree to the bootstrap sample until the minimum node size is reached by:
   (a) Selecting m variables at random
   (b) Pick the best split and split the node
3. Repeat until you have a forest

The classification is then a majority vote in the forest.

4.9.5 Support Vector Machines

Support Vector Machines (SVM) are a powerful type of classifiers. They are are a generalisation of support vector classifiers, which are again a generalisation of optimal separating hyperplanes. This review is based on [Hastie et al., 2009].

The optimal separating hyperplane, finds the optimal linear separation of two non-overlapping classes and maximises the distance to the closest point from each class.

Let us define a hyperplane by

$$f(x) = x^{\mathrm{T}}\beta + \beta_0 = 0 \tag{4.66}$$

where β is a unit vector and we classify based on the sign of $f(x)$. Our training data are N pairs of (x, y) where x is a measure and y is the classification (-1,1).

The hyperplane can be found by maximising the margin $M = 1/\|\beta\|$:

$$\begin{aligned} \underset{\beta,\beta_0}{\text{maximise}} \quad & M \\ \text{subject to} \quad & y_i(x_i^{\mathrm{T}}\beta + \beta_0) \geq M, \quad i = 1, \cdots, N \\ & \|\beta\| = 1 \end{aligned}$$

or more conveniently formulated:

$$\begin{aligned} \underset{\beta,\beta_0}{\text{minimise}} \quad & \|\beta\| \\ \text{subject to} \quad & y_i(x_i^{\mathrm{T}}\beta + \beta_0) \geq 1, \quad i = 1, \cdots, N \end{aligned}$$

The support vector classifier extends this to the case where the classes overlap by modifying the constraint:

$$y_i(x_i^{\mathrm{T}}\beta + \beta_0) \geq M(1 - \zeta_i) \tag{4.67}$$

where $\zeta > 0$ is a so called slack variable. Without going into too much detail, this can be formulated as a convex problem of which the dual can be solved efficiently by maximising:

$$L_D = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{i'=1}^{N} \alpha_i \alpha_{i'} y_i y_{i'} x_i^{\mathrm{T}} x_{i'} \tag{4.68}$$

subject to a number of constraint (see [Hastie et al., 2009] p. 420).

The extension to SVM is now simple. We choose a suitable basis expansion $h(x)$ for x . The beauty is, that they only appear through their inner product in the dual function:

$$L_D = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{i'=1}^{N} \alpha_i \alpha_{i'} y_i y_{i'} \langle h(x_i), h(x_{i'}) \rangle \tag{4.69}$$

and the decision border is now:

$$f(x) = h(x)^{\mathrm{T}}\beta + \beta_0 \tag{4.70}$$

$$= \sum_{i=1}^{N} \alpha_i y_i \langle h(x), h(x_i) \rangle + \beta_0 \tag{4.71}$$

4.9.6 Sparse methods

It should be mentioned, that in addition to those methods already mentioned, a large number of sparse methods exist e.g. lasso, least angle regression [Hastie et al., 2009], and Sparse discriminant analysis [Clemmensen et al., 2011]. As more classical methods worked adequately on the problems in this thesis they are, however, outside the scope here.

Part II

Scientific Contributions

Chapter 5

NanoGuide project: Paper A, B, and C

One of the goals of the NanoGuide project was to develop gel based injectable markers for use in radiotherapy. It was successful in that regard, see paper A and B. The gel described in paper A was further developed for use in in-vivo dosimetry, see paper C.

5.1 Paper A

Paper A presents a gel-based x-ray marker based on gold nano-particles. It was demonstrated that the marker had a high contrast, was stable and biocompatible. The gel was tested in both mice for stability and biocompatibility and in a dog with a spontaneous tumour as proof of concept for guidance in radiotherapy. The mice experiment consisted of repeated measurements on 5 mice using a micro-CT scanner. My task was analysis of stability and homogeneity of the gel.

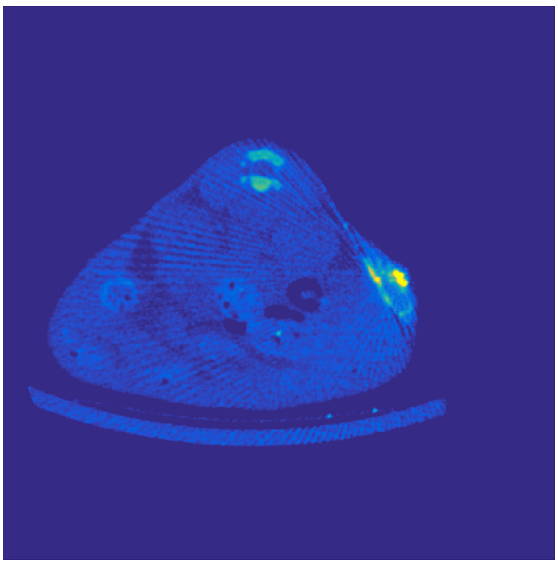


Figure 5.1: A slice in a mouse showing the Au-nanoparticle gel. The gel is the bright structure to the right.

5.1.1 Methods and Results

The gels were quite inhomogeneous and a manual segmentation was necessary. Fig. 5.1 should help illustrate why. The volume showed no major changes once the ethanol that makes the gel liquid had diffused out. By inspecting the cumulative histograms it was clear that the inhomogeneity increased over time, see fig. 5.2. To test if the gels were more inhomogeneous than an earlier formulation with less contrast [Jølck et al., 2014], a common Box-Cox transformation (sec. 4.6.1) for the two groups was performed to obtain approximate normality followed by a Wilcoxon rank-sum test (sec. 4.5.1) of the variances of the two groups. The variances of the gels were found not to differ significantly (P=0.0734). Thus, even with more contrast the new formulation was not more inhomogeneous than the old.

5.1.2 Contributions

- Segmenting gels
- Presented data visually to establish that the volume and overall contrast was stable and the heterogeneity of gel increased

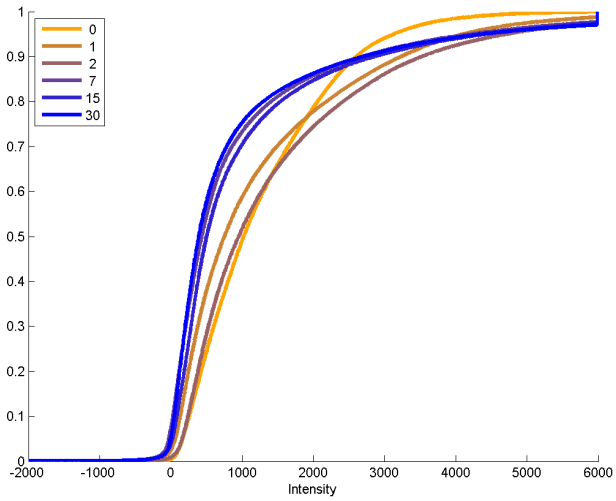


Figure 5.2: Cumulative histogram from a representative mouse. The colours represent different days.

- Statistical analysis of inhomogeneity compared to older formulation

5.2 Paper B

Although the gel described in paper A, worked well, an even better formulation was found. An in-vivo test of the degradation rate of this gel is described in paper B. By using iodine instead of gold particles, the contrast level could be increased and the homogeneity and stability were even better. With regards to production, storage and use the gel also had favourable qualities. This paper investigated the degradation rate of two groups with different contrast levels.

5.2.1 Methods and Results

As the gels were much more homogeneous it was easy to segment them using a Matlab implementation of Chan-Vese (sec. 3.1.2). This performed very well, see fig. 5.3. To analyse both group difference and degradation a repeated measurements model (sec. 4.2) was used. The data from the two groups are shown

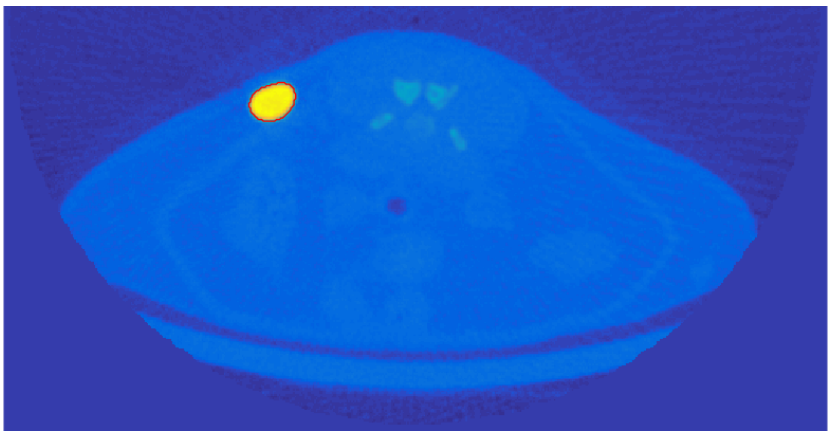


Figure 5.3: A slice in a mouse showing the iodine gel. The red curve is the segmentation using Chan-Vese

in fig. 5.4. A clear downwards trend is seen, but apart from that it is difficult to make conclusions.

A number of models were tested, to see which one most adequately described the data. k is the group, j is the mouse and i is the observation. The errors were

Table 5.1

Model overview		
M1	$\alpha(k)_j + \beta(k)_j\, t(k)_{ij}$	Individual intercept and slope
M2	$\alpha(k)_j + \beta(k)\, t(k)_{ij}$	Individual intercept and group slope
M3	$\alpha(k)_j + \beta\, t(k)_{ij}$	Individual intercept and common slope
M4	$\alpha(k)_j$	Individual intercept
M5	$\alpha(k)$	Group intercept

modelled as autoregressive, giving the following variance-covariance matrix for a single mouse:

$$D(\epsilon_{jk}) = \sigma^2 \begin{bmatrix} 1 & \cdots & \rho^{|t(k)_{1j} - t(k)_{Ij}|} \\ \vdots & \ddots & \vdots \\ \rho^{|t(k)_{1j} - t(k)_{Ij}|} & \cdots & 1 \end{bmatrix} \tag{5.1}$$

where I is the last observation.

Using Akaikes information criterion corrected for finite sample size (sec. 4.4.2.1) M3 found to be an appropriate model, i.e. individual intercept and common

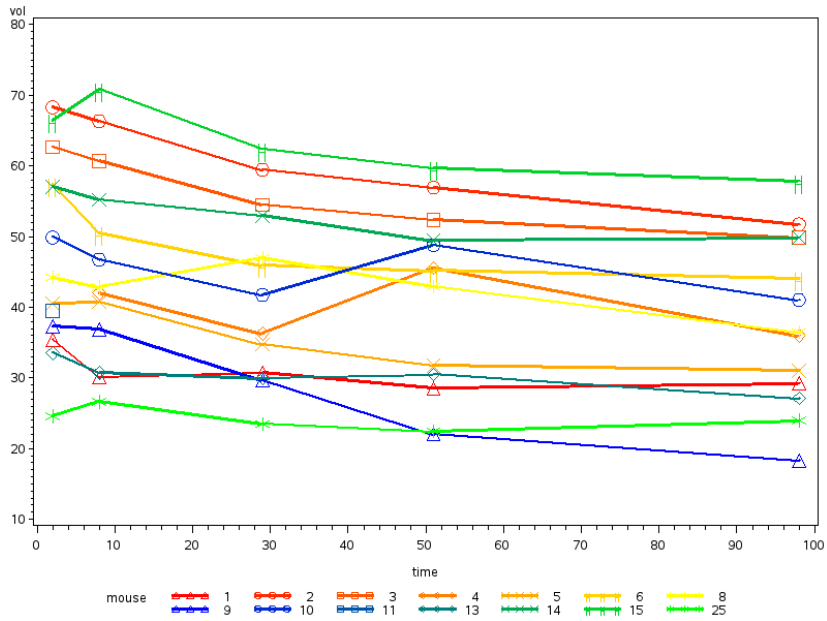


Figure 5.4: Volume data over time from mice. 1-8 is one group and 9-25 another group.

slope. Thus concluding that the two groups did not differ in degradation rate. The degradation of the gels were simulated with a simple physical model assuming that the degradation only occurred in the surface. With this model an approximate time to total degradation could be made.

5.2.2 Contributions

- Segmenting gels
- Statistical analysis of difference in two formulations
- Statistical analysis of degradation rate
- Simple physical model for the degradation of the gels

5.3 Paper C

Returning to the first gel (paper A), I got an idea to use it in in-vivo dosimetry. Based on this paper [Hansen et al., 2008], the gold nano-particles could be exchanged with silver. When irradiated with high-energy photons a (γ,n)-reaction can occur and the silver can become β^+ active (sec. 2.6). The amount of radioactivity can then be measured in a PET scanner and the dose calculated on basis of the activity.

5.3.1 Methods and Results

Gels with silver were prepared and stored in plastic cylinders filled with gelatine. They were inserted into a thorax phantom, and irradiated at three different dose-levels. 2 Gy, 10 Gy, and 22 Gy. Immediately afterwards they were transported to a PET-scanner and scanned for half an hour.

The scannings were reconstructed using a Filtered Back-Projection (FBP). While generally more noisy than iterative methods it does not suffer from bias in low activity regions (sec. 2.4.1). The activities of the gels were manually extracted by centring a ball with radius 20 mm on each gel and taking the mean value.

We then wanted to establish whether the position, the individual gels and the dose had any influence on the measured activity. Obviously we had an expectation that the dose might had something to say. The first model were with all

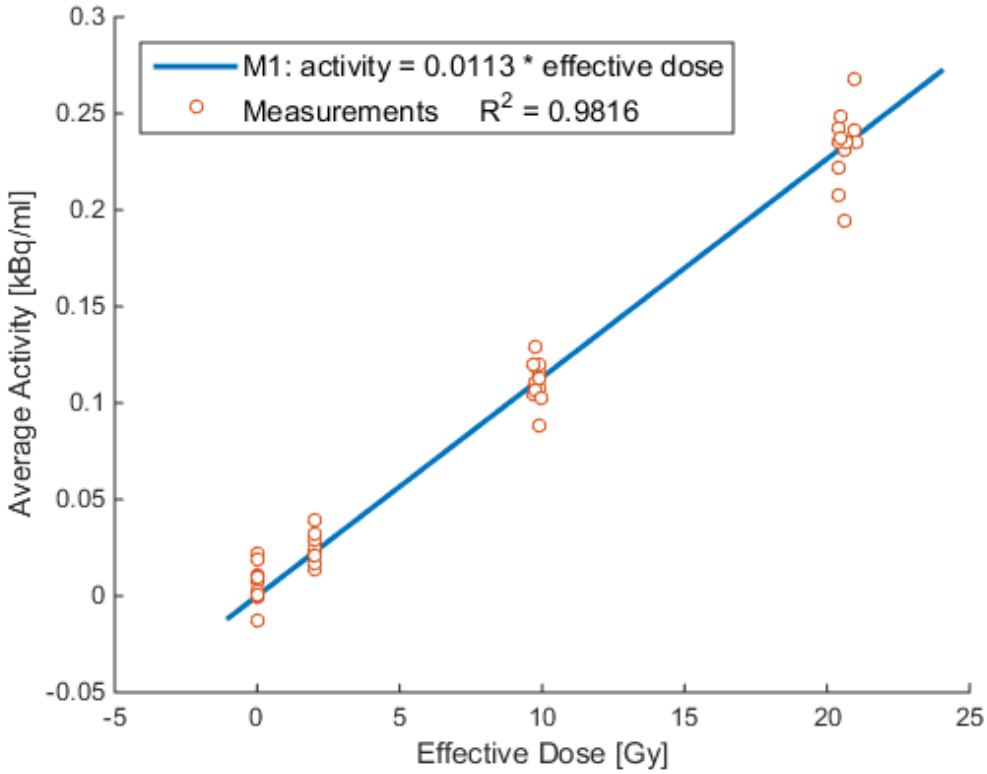


Figure 5.5: The activities of the gels plotted against the effective dose

effects and their pairwise interaction. Then all the effects, then the dose and intersect and finally only the dose.

By successive testing we showed that the data could be described using a simple proportionality model, see fig. 5.5.

5.3.2 Contributions

- Conceiving experiment
- Performing PET/CT scan
- Extracting activity from scans

- Statistical analysis of effects

Chapter 6

Paper D

In this paper the data were MRI-scans of the brains of alzheimers patients. The hypothesis was that exercise delayed the progression of alzheimers disease. For that purpose the subjects were divided into an intervention group that exercised and a control group that did not. All subjects were scanned twice with 16 weeks in between and the brains were segmented using FreeSurfer software [Reuter et al., 2012, Fischl et al., 2002]. The resulting segmentation gave volumetric measures and the thickness of various gyri and sulci.

6.1 Methods and Results

To cancel within subject correlations, the relative change from pre-scan to post scan were used. Age and sex were used as covariates in all tests. The group differences were investigated using Hotellings T^2-test (sec. 4.1.5). As outliers were present the results were validated by a non-parametric method, the Oja rank test (sec. 4.5.2).

Several cognitive measures were investigated through tests. It was hypothesised, that they each correlated with specific brain regions. To investigate a potential correlation of a single test score with multiple brain regions we tested the co-

variance matrix (sec. 4.7). If we have a test score $\mathbf{x}$ and a brain region with two subfields $\mathbf{y_1}, \mathbf{y_2}$ the null-hypothesis was:

$$H_0 : \begin{bmatrix} \mathrm{cov}(\mathbf{x},\mathbf{x}) & \mathrm{cov}(\mathbf{x},\mathbf{y_1}) & \mathrm{cov}(\mathbf{x},\mathbf{y_2}) \\ \mathrm{cov}(\mathbf{y_1},\mathbf{x}) & \mathrm{cov}(\mathbf{y_1},\mathbf{y_1}) & \mathrm{cov}(\mathbf{y_1},\mathbf{y_2}) \\ \mathrm{cov}(\mathbf{y_2},\mathbf{x}) & \mathrm{cov}(\mathbf{y_2},\mathbf{y_1}) & \mathrm{cov}(\mathbf{y_2},\mathbf{y_2}) \end{bmatrix} = \begin{bmatrix} \mathrm{cov}(\mathbf{x},\mathbf{x}) & 0 & 0 \\ 0 & \mathrm{cov}(\mathbf{y_1},\mathbf{y_1}) & \mathrm{cov}(\mathbf{y_1},\mathbf{y_2}) \\ 0 & \mathrm{cov}(\mathbf{y_2},\mathbf{y_1}) & \mathrm{cov}(\mathbf{y_2},\mathbf{y_2}) \end{bmatrix} \quad (6.1)$$

against all alternatives. If we rejected the null-hypothesis a post-hoc test of the individual correlations was performed.

6.2 Contributions

- Statistical analysis of difference in groups - both parametric and non-parametric
- Statistical analysis of correlation between cognitive test-scores and brain-regions

Chapter 7

Paper E

The distribution of abdominal fat has been linked to various metabolic risk factors and is of great clinical interest. We wanted to make an integrated software solution that could segment a MRI scan fully automatically. Further it should be robust to changes in sequence and across different population groups.

7.1 Methods and Results

The scans were bias-corrected using a generative model (sec. 3.4). By formulating the segmentation as a graph-cut problem (sec. 3.1.3), we can utilise the information in adjacent slices and achieve a globally 'optimal' segmentation of the subcutaneous fat. The visceral fat was segmented using the inner surface of the subcutaneous fat as a mask. A dynamic threshold based on k-means clustering was used.

A set of settings for the bias-correction and energies for the graph-cut problem was found, that performed robustly. These energies consist of a weighting of various gradient and surface based costs. In the same framework a method for exclusion of the non-visceral fat around the spine was implemented. When evaluated against an expert observer the method performed very well. An example of the segmentation of subcutaneous fat can be seen in fig. 7.1.

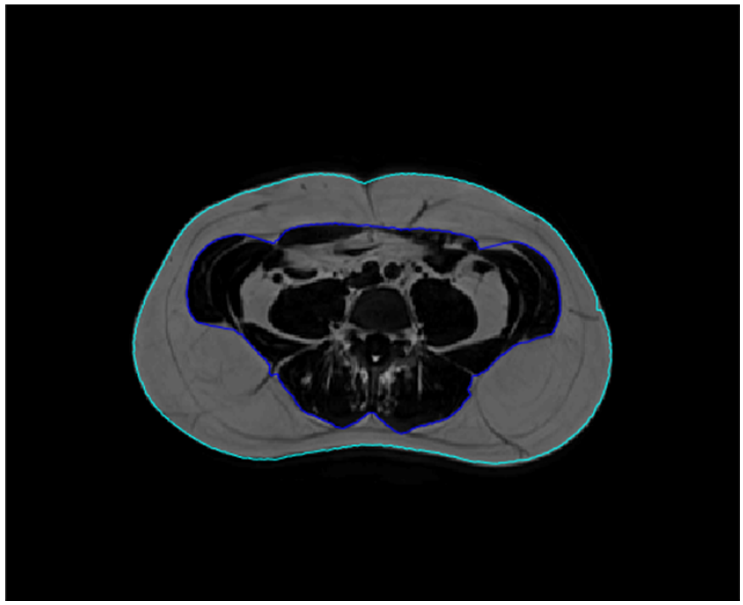


Figure 7.1: An example of segmentation of the subcutaneous fat

7.2 Contributions

- Integrated bias-correction and graph-cut in the pipeline
- Found energies and settings that worked robustly
- Method for spine extraction
- Demonstrated usefulness in three different studies with different sequences and population segments

Chapter 8

Paper F

Phase-contrast imaging is an emerging imaging modality for medical use. It allows for imaging of absorption, dark-field and differential phase. The dark-field in particular shows a superior sensitivity to changes in the lungs. While both emphysema and fibrosis can be detected using the dark-field, it has been questioned whether they can be separated as they exhibit the same characteristics. The focus of paper F was to investigate this and if possible classify between healthy, emphysematous, and fibrotic lung tissue. In addition we investigated if the differential phase image had information that could contribute to differentiation.

8.1 Methods and Results

By using an Active Appearance Model (AAM) (sec. 3.1.1) the lungs were segmented and deformed to a standard shape. This allowed for comparisons across different mice. A medical doctor contoured diseased region for use in a training and test set. Methods for automatic extraction of the ribs and spine were implemented. By excluding the ribs and spine, a model of the healthy lung was generated by averaging a set of healthy mice. By normalising with the healthy lung model we could incorporate regional information.

With regards to the differential phase we tested three texture filters (sec. 3.2.2) as a mean to extract information: The range, standard deviation, and entropy filter. Each with various diameter kernels. The marginal probabilities for the three classes (healthy, emphysematous, and fibrotic) were estimated for absorption, dark-field, differential phase, and the filters applied to the differential phase. Using the healthy lung-model, a Naive Bayes classifier (sec. 4.9.1) could differentiate and classify with an accuracy of 84.05 %. An example can be seen in fig. 8.1.

8.2 Contributions

- Method for rib and spine removal
- Implemented classification algorithms and performed all analysis in that regard
- Statistical analysis of the results

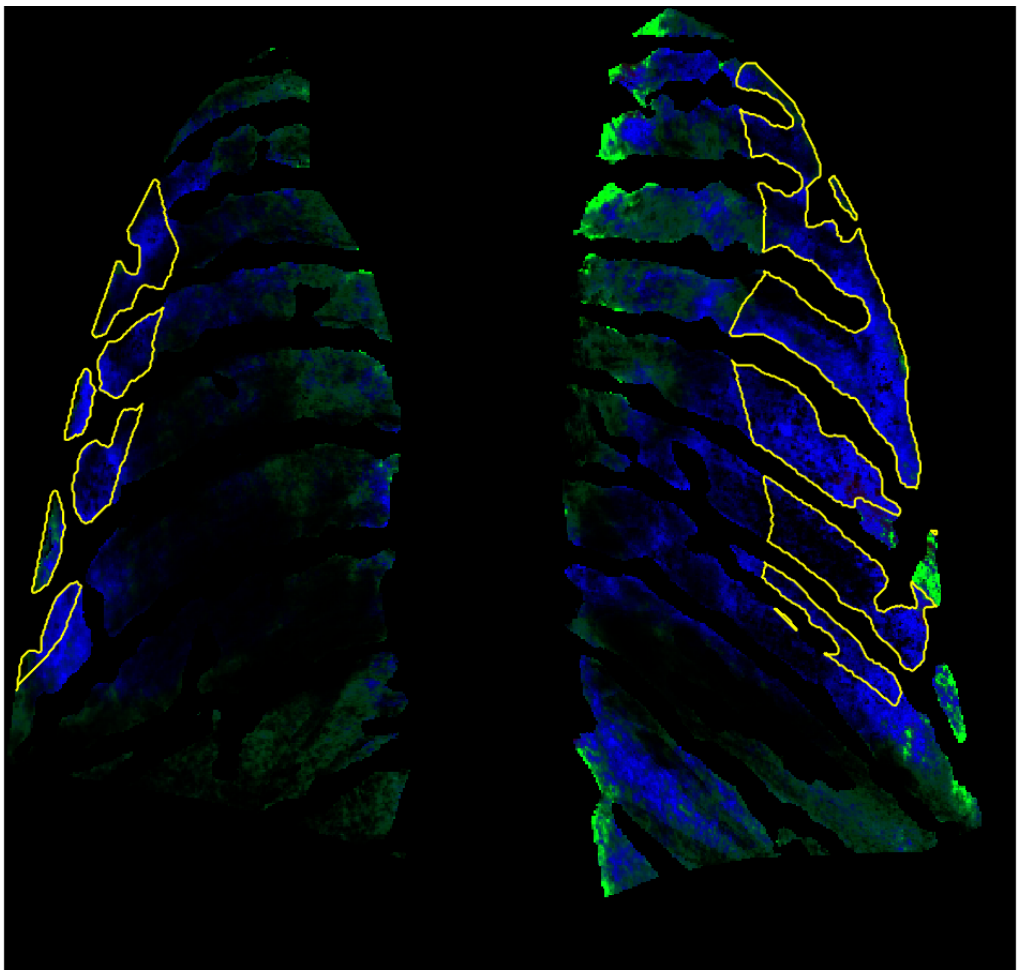


Figure 8.1: An emphysematous lung showing good agreement with classification. Green: healthy. Blue: Emphysema. Red: fibrosis. The yellow line is the region demarcated as emphysema by an expert.

Chapter 9

Conclusion

As demonstrated in this thesis data analysis of medical images is a wide topic with many problems. Here a small subset of them have been tackled. In all the papers contributing to this thesis, segmentation was the first step in the analysis if not the end-goal. When having extracted data through a segmentation or other means, statistical analysis is often required.

In the NanoGuide project two gels were developed to function as x-ray markers for radiotherapy. The necessary data analysis for both publications and regulatory approval was performed. This demonstrates the need for both extraction data, though segmentation, and rigorous analysis through the use of statistics. The statistical analysis was used to confirm that gold-nanoparticle gels with different contrast levels did not differ in inhomogeneity. For the iodine based gels with different contrast levels it was confirmed that the degradation rate did not differ through the use of a RMM.

The gel with gold-nanoparticles was tested for a new application in in-vivo dosimetry with the gold exchanged with silver. This use was successfully demonstrated in a thorax-phantom. It was shown that neither position nor inter-gel variance had any practical effect on the results.

The analysis of MRI brain scans of alzheimer patients also started with a segmentation. Statistical analysis was applied to make conclusion about group

differences and correlation between cognitive test-scores and changes in brain regions.

A software package in development for segmentation of abdominal was presented. The method works robustly across different sequences and population groups.

The efficacy of phase-contrast x-rays in differentiation different lung diseases was demonstrated. In a murine model we successfully classified between healthy, emphysematous, and fibrotic tissue.

Through solutions to this diverse set of problems the objectives of the thesis was met.

Part III

Included Papers

Appendix A

Injectable Colloidal Gold for Use in Intrafractional 2D Image-Guided Radiation Therapy

ADVANCED HEALTHCARE MATERIALS
www.advhealthmat.de

Materials Views
www.MaterialsViews.com

FULL PAPER

Injectable Colloidal Gold for Use in Intrafractional 2D Image-Guided Radiation Therapy

*Rasmus I. Jølck, Jonas S. Rydhög, Anders N. Christensen, Anders E. Hansen, Linda M. Bruun, Henrik Schaarup-Jensen, Asger Stevner von Wenck, Betina Børresen, Annemarie T. Kristensen, Mads H. Clausen, Andreas Kjær, Knut Conradsen, Rasmus Larsen, Per Munck af Rosenschöld, and Thomas L. Andresen**

In the western world, approximately 50% of all cancer patients receive radiotherapy alone or in combination with surgery or chemotherapy. Image-guided radiotherapy (IGRT) has in recent years been introduced to enhance precision of the delivery of radiation dose to tumor tissue. Fiducial markers are often inserted inside the tumor to improve IGRT precision and to enable monitoring of the tumor position during radiation therapy. In the present article, a liquid fiducial tissue marker is presented, which can be injected into tumor tissue using thin and flexible needles. The liquid fiducial has high radiopacity, which allows for marker-based image guidance in 2D and 3D X-ray imaging during radiation therapy. This is achieved by surface-engineering gold nanoparticles to be highly compatible with a carbohydrate-based gelation matrix. The new fiducial marker is investigated in mice where they are highly biocompatible and stable after implantation. To investigate the clinical potential, a study is conducted in a canine cancer patient with spontaneous developed solid tumor in which the marker is successfully injected and used to align and image-guide radiation treatment of the canine patient. It is concluded that the new fiducial marker has highly interesting properties that warrant investigations in cancer patients.

1. Introduction

Nanomaterials offer unique properties that can provide society with new healthcare technologies for treatment of disease. With the increasing aging of the population in the western world, cancer incidence continues to rise creating a crucial demand for efficient and cost-effective therapies. Radiation therapy is a cost-effective treatment modality and approximately 50% of cancer patients receive radiation therapy after being diagnosed with cancer.[1] Modern radiation oncology relies on advanced, high-precision imaging techniques such as computed tomography (CT), positron emission tomography (PET), and magnetic resonance (MR) to enable the delivery of high radiation doses to defined targets.[2–4] However, tumors rarely display a fixed position during irradiation or within the treatment period

Dr. R. I. Jølck, Dr. A. E. Hansen, L. M. Bruun, Prof. T. L. Andresen
DTU Nanotech, Department of Micro-and Nanotechnology
Center for Nanomedicine and Theranostics
Technical University of Denmark
Building 345E, Ørsteds Plads,
2800 Lyngby, Denmark
E-mail: thomas.andresen@nanotech.dtu.dk
J. S. Rydhög, Dr. P. Munck af Rosenschöld
Department of Oncology
Section of Radiotherapy
3994, Rigshospitalet
Blegdamsvej 9, 2100 Copenhagen, Denmark
J. S. Rydhög, Dr. P. Munck af Rosenschöld
the Niels Bohr Institute, University of Copenhagen
Blegdamsvej 17, 2100 Copenhagen, Denmark
A. N. Christensen, Prof. K. Conradsen, Prof. R. Larsen
DTU Compute, Department of Applied Mathematics and Computer Science
Technical University of Denmark
Building 321/324, Matematiktorvet, 2800 Lyngby, Denmark
H. Schaarup-Jensen, Prof. M. H. Clausen
DTU Chemistry
Department of Chemistry
Center for Nanomedicine and Theranostics
Technical University of Denmark
Building 207, Kemitorvet, 2800 Lyngby, Denmark
A. Stevner von Wenck, B. Børresen, Prof. A. T. Kristensen
Department of Veterinary Clinical and Animal Sciences
Faculty of Health and Medical Sciences
University of Copenhagen
Dyrlægevej 16, 1870 Frederiksberg C, Denmark
Prof. A. Kjær
Department of Clinical Physiology
Nuclear Medicine & PET and Cluster for Molecular Imaging
Rigshospitalet and University of Copenhagen
Blegdamsvej 9, 2100 Copenhagen, Denmark

DOI: 10.1002/adhm.201400651

due to breathing motion, changes in organ filling, and tumor size.[5–7] To compensate for variations in tumor position frequent imaging (kilovoltage radiographs or fluoroscopy—also referred to as 2D X-ray imaging) during the course of treatment or image-guided radiation therapy (IGRT) is now being used in the clinic to improve radiation precision and accuracy, thereby reducing the treatment toxicity and optimizing therapeutic efficacy.[2,6,8] When the tumor position cannot be correlated with internal reference points such as the skeleton, tumor localization can be enhanced by alignment to radiographic markers implanted in or near the tumor.[9–11] Such markers can improve treatment through optimized planning, tumor positioning accuracy, and tumor tracking resulting in more precise dose delivery and enable treatment of cancers, which are presently not eligible for radiotherapy.[12,13] Tissue markers currently used in the clinic are metal-based solid implants with large physical dimensions, which require complicated insertion procedures and therefore the risk of complications. Lung tumor insertion is especially problematic, with pneumothorax (reported in 33%–68% of patients) as well as bleeding being reported as primary complications.[14–16] In order to increase the usability and patient compliance of radiographic markers, we recently described the concept of a liquid injectable fiducial marker by exploiting polyethylene glycol (PEG)-coated gold nanoparticles (AuNPs) in a nonsolvent-induced phase separation (NIPS) gelation matrix.[17] This system was found to be biocompatible and clearly visible on 3D X-ray imaging, However, due to the chemical properties of the encapsulated AuNPs only 10 mg AuNPs mL^{-1} providing approximately 200 Hounsfield units (HU) of contrast could be included, which is insufficient for 2D X-ray visualization. For gating-based radiation procedures, often exploited in for example lung cancer patients, 2D X-ray visualization is required for intrafractional real-time imaging of moving tumors.[6,18–20] Additionally, fiducial marker visualization using 2D X-ray techniques lowers the X-ray exposure level to the patients and may reduce treatment time compared to 3D X-ray-based procedures.[2] Furthermore, with the increasing focus on stereotactic radiation procedures and the clinical introduction of proton therapy, visualization of tumor position and motion becomes crucial to optimize irradiation of cancerous tissue.[21,22] Therefore, radiopaque fiducial markers continue to be in high demand and development of new marker systems is of paramount importance.

In the present work, we have investigated a novel liquid fiducial marker (nanogel) system based on sucrose acetate isobutyrate (SAIB) and uniform poly(*N*-isopropyl acrylamide) (PNIPAM)-coated AuNPs with radio-opacity adequate for 2D X-ray visualization. SAIB, a fully acylated derivative of sucrose, is a biodegradable nonpolymeric compound, which has been approved for oral consumption and is generally regarded as safe (GRAS) by the FDA. SAIB mixtures with ethanol (EtOH) behaves as Newtonian liquids and have a low viscosity (50–200 mPa s) suited for injection using thin hypodermic needles.[23–25] Upon hydration, an amorphous (gel-like) solid is formed, which efficiently encapsulate high concentrations (>100 mg Au mL^{-1}) of PNIPAM-coated AuNPs providing contrast levels sufficient for visualization in 2D X-ray imaging techniques such as fluoroscopy and ExacTrac imaging (Brainlab AG, Feldkirchen, Germany). The developed nanogel has been characterized in vitro and the in vivo stability and visibility was evaluated in immunocompetent mice over a period of 8 weeks. Proof-of-concept in a clinical canine cancer patient with a spontaneous growing tumor was additionally established and the developed nanogel with PNIPAM-coated AuNPs was applied in IGRT in a canine cancer patient where marker visibility, stability, and marker migration were evaluated over time during the course of radiotherapy. The developed nanogel fulfills the requirements for use in guidance of radiation therapy as it is easy to inject, remains stable over the full time course of radiation treatment (6 weeks), shows no sign of migration, and is the first liquid fiducial marker to enable visualization using both 2D and 3D X-ray-based imaging techniques.

2. Results and Discussion

Visualization of fiducial markers in 2D X-ray imaging requires high electron density within the nanogel to possess efficient X-ray attenuation/radio-opacity, which can be achieved by doping the injectable gel with colloid AuNPs. We recently exploited the use of PEG-coated AuNPs but due to the high hydrophilicity of the PEG-coated AuNPs only 10 mg AuNPs mL^{-1} could be included in the liquid marker, which was insufficient for 2D X-ray visibility. Higher AuNP-concentrations resulted in a significant burst release due to the higher preference of the hydrophilic PEG-coated AuNPs for the surrounding aqueous phase.[17]

2.1. Engineering AuNPs for Enhanced Nanogel Compatibility

In order to increase the concentration of AuNPs and enhance the compatibility of the AuNPs with the hydrophobic SAIB-based gelation matrix, two different strategies were investigated: i) Masking the AuNPs by chemisorption of a dithiolane SAIB derivate ($(SH)_2$-SAIB) (**5**) (see Section 2.1.1, and Supporting Information) and ii) using commercially available thiol-functionalized PNIPAM with a molecular weight of 3.500 Da ($SH\text{-}PNIPAM_{3.500}$). Both strategies resulted in hydrophobic AuNPs, which was anticipated to reduce the previously observed AuNP burst release from the nanogels. Hence, we hypothesized that realization of one or both of these strategies should enable formulation of nanogels with higher concentrations of AuNPs resulting in higher X-ray contrast levels.

2.1.1. Synthesis of SAIB-Based Capping Agent for AuNP Functionalization

The dithiolane-functionalized SAIB derivate **5** was prepared as outlined in Scheme S1 (Supporting Information) starting from sucrose (**1**). Mono silyl protection of the 6′-OH position of the fructofuranose unit was achieved using *tert*-butyl(chloro)diphenylsilane (TBDPS-Cl) and 4-dimethylaminopyridine (DMAP) yielding 6′-O-TBDPS-sucrose (**2**).[26] The remaining hydroxyl groups were isobutyrylated using isobutyric anhydride and DMAP resulting hepta-isobutyryl-6′-O-TBDPS-sucrose (**3**). Initial attempts of deprotection of the TBDPS group using

ADVANCED HEALTHCARE MATERIALS
www.advhealthmat.de
www.MaterialsViews.com

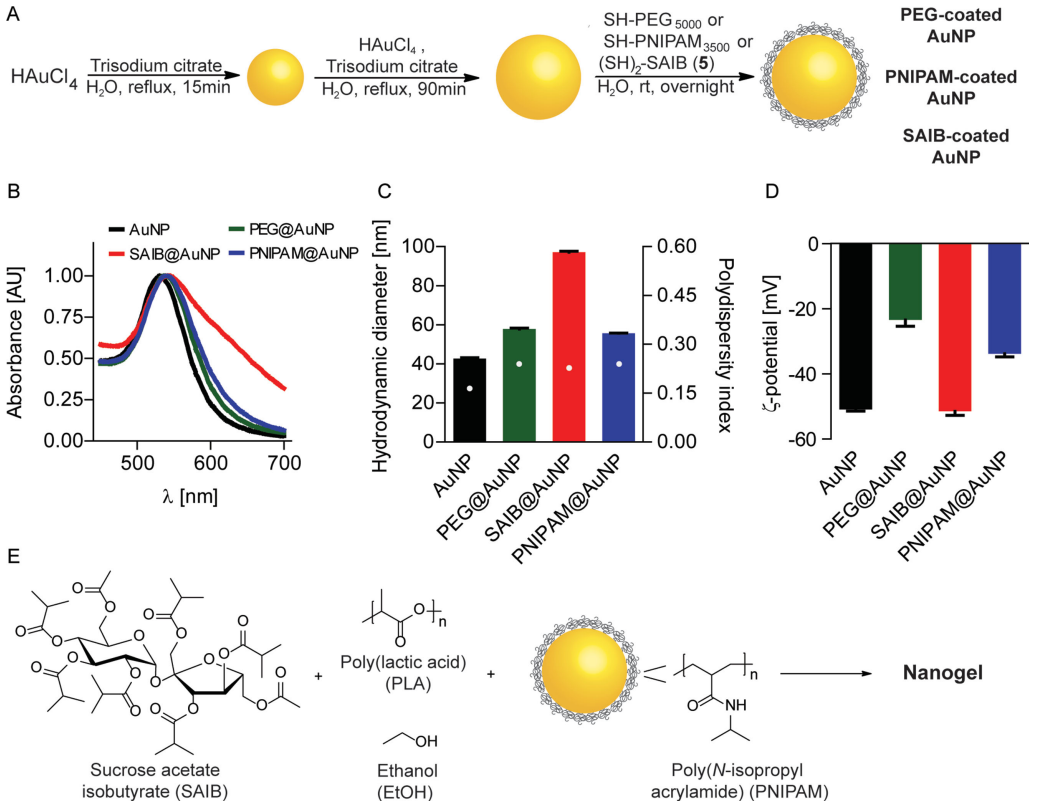


Figure 1. A) Synthesis protocol for obtaining PEG-, PNIPAM, and SAIB-coated AuNPs based on a three steep seeding protocol using chloroauric acid as Au^{3+} source and trisodium citrate as reductant and stabilizer; B) UV–vis characterization of PEG-, PNIPAM-, and SAIB-coated AuNPs in aqueous solution; C) DLS characterization of PEG-, PNIPAM-, and SAIB-coated AuNPs in aqueous solution; and D) ζ-potential characterization of PEG-, PNIPAM-, and SAIB-coated AuNPs in aqueous solution. E) Nanogel composition and preparation procedure.

tetra-n-butylammonium fluoride (TBAF) indicated a high degree of isobutyryl migration; fortunately, this could be suppressed by addition of acetic acid leading to hepta-isobutyryl-sucrose (**4**). Finally, the dithiolane moiety was introduced by *N*-(3-dimethylaminopropyl)-*N*′-ethylcarbodiimide (EDC) activation of lipoic acid in the presence of DMAP yielding the dithiolane-functionalized SAIB derivate **5** (a detailed description of synthesis and compound characterization is available as Supporting Information).

2.1.2. AuNP Synthesis, Functionalization, and Characterization

Citrate-stabilized AuNPs were synthesized using a seeding procedure as previously described resulting in uniform electrostatic-stabilized AuNPs with a diameter of approximately 40 nm.[17] Subsequently, chemisorption of SH-PEG_{5000}, SH-$PNIPAM_{3500}$ or $(SH)_2$-SAIB (**5**) to the citrate-stabilized AuNPs was conducted in aqueous solution in order to increase the colloid compatibility with SAIB-based gels (**Figure 1**A). Increasing the compatibility of the AuNPs with the SAIB-matrix includes increasing the hydrophobicity of the particles. However, a delicate balance exists where the hydrophobicity of the particles is increased without compromising their stability in aqueous solution during synthesis. As evident from the UV–vis (Figure 1B) and dynamic light scattering (DLS) (Figure 1C) analysis, AuNP coating with PEG and PNIPAM was successful. Only a slight increase in the surface plasmon resonance was observed, which is expected due to the microenvironmental changes on the surface of the particles. Additionally, DLS confirmed the presence of the two polymers at the colloid interface since an increase of 15.1 and 12.9 nm was observed for the PEG-coated and PNIPAM-coated AuNPs, respectively. However, chemisorption of $(SH)_2$-SAIB (**5**) to the surface of the electrostatic-stabilized AuNPs caused considerably aggregation due to the poor aqueous solubility of the formed particles, even under the very diluted conditions used. The aggregates formed from SAIB-coated AuNPs were isolated but redispersion in organic solvents such as absolute ethanol was not possible; thus, this class of particles was not investigated further. Synthesis and

subsequently coating of AuNPs with $(SH)_2$-SAIB (**5**) in pure organic solvent using the Burst–Shiffin method may be feasible, but this was not attempted due to the small size of particles formed by this method (<10 nm) and the fact that non-biocompatible solvents such as toluene and toxic phase-transfer agents have to be used, which could limit their use in biological settings.[27]

Additionally, handling of PNIPAM-coated AuNPs was found to be superior compared to PEG-coated AuNPs as these could be lyophilized and stored for prolonged time as an air stabile nanopowder, which readily dispersed into MQ-H_2O and a wide range of protic- and aprotic organic solvents including anhydrous ethanol without causing any changes to particle morphology (Figure S1, Supporting Information). This enabled easy up-concentration, storage, and nanogel formulation compared to the previous described method where AuNPs were transferred into anhydrous EtOH by centrifugation prior to use.

2.2. Contrast Evaluation for 2D X-Ray Visualization of Nanogel in Humans

In order to evaluate the required concentration of AuNPs within the nanogel for visibility on 2D X-rays images in humans, a simulation of 2D digitally reconstructed radiographs (DRRs) was performed based on clinical 3D CT-data for a patient diagnosed with lung cancer. A lung tumor model was chosen as this region is considered as a worst-case scenario due to the high concentration of bone structure in the thoracic region. A simulated nanogel (500 µL) was digitally placed inside the lung tumor of the patient and the specified HU contrast level was varied from 500 to 2000 HU with increments steps of 500 HU on the 3D CT-data. 2D DRRs were created for all angles entering perpendicular around the patient in steps of 5° using a ray-tracing algorithm[28,29] resulting in 72 simulated projections for each contrast level (**Figure 2**). The center position of the DRRs was chosen to be in the center of the simulated nanogel placed centrally within the lung tumor of the patient. On the resulting DRRs, the visibility for each contrast level was evaluated by a contrast-to-noise ratio (CNR) approach, where a region of interest (ROI) for one projection and contrast level was contoured to define the nanogel and then transferred to all other projections. The mean intensity (M_m) and standard deviation (σ_m) of ROIs were calculated and compared to the mean (M_b) and standard deviation (σ_b) from a ROI on a DRR with no simulated nanogel present according to a CNR evaluation (Equation (1)).

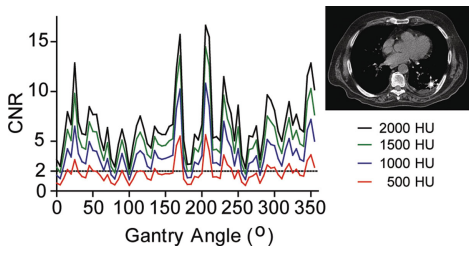


Figure 2. Contrast to noise (CNR) level versus angle of entrance for a simulated nanogel (500 µL) with various contrast levels (500–2000 HU) placed digitally inside a lung tumor.

$$\mathrm{CNR} = \frac{|M_m - M_b|}{\sqrt{\sigma_m^2 - \sigma_b^2}} \quad (1)$$

The CNR level can be considered a quantitative measurement of the visibility of the object measured compared to the background contrast from the surroundings. Generally, high CNR translate to increased visibility of the marker. We considered the simulated nanogel visible in the 2D DRR if CNR > 2. This value corresponds to a significance level of 0.05 obtained from a two-tailed *t*-test with a large number of degrees-of-freedom.[30] As evident from Figure 2, some angles requires a higher X-ray contrast level of the nanogel in order to obtain a CNR > 2 as these penetrates bone structure prior to reaching the nanogel. Other angles are more easily accessible and thereby result in lower background contrast, thus making the nanogel more visible on the 2D DRRs. Based on the lung cancer patient projection simulation on the DRRs, it was evident that the contrast level required for CNR > 2 in 95% of all perpendicular angles was approximately 1000 HU (spherical nanogel with a total volume of 500 µL) corresponding to a nanogel with approximately 30 mg AuNPs mL^{-1}. A comparison of the 2D X-ray contrast level from PNIPAM-coated AuNPs (10, 30, 60, and 90 mg mL^{-1}) and clinically used solid fiducial markers is provided in Figure S2, Supporting Information.

2.3. In Vitro Evaluation of Nanogel Performance

In vitro release kinetics of encapsulated PEG-coated and PNIPAM-coated AuNPs in SAIB nanogels doped with polylactic acid (PLA) with the composition SAIB/EtOH/PLA (75:20:5) + 30 mg AuNPs mL^{-1} were examined by placing small droplets (200 µL) of nanogels into MQ-H_2O at physiological temperature. Small aliquots of MQ-H_2O were removed at specific time intervals and replaced with fresh MQ-H_2O to mimic the sink effect in vivo. The amount of released AuNPs was determined by UV–vis spectroscopy using a standard curve of individual synthesized and coated AuNPs with a known concentration (Figure S3, Supporting Information). Injection of nanogel into aqueous solution causes an efflux of EtOH from the nanogel into the aqueous compartment. Efflux of EtOH results in rapid increase in gel viscosity,[24] but may also induce a burst release of encapsulated AuNPs.[24] Indeed, a rapid burst release of PEG-coated AuNPs from nanogels composed of SAIB/EtOH/PLA (75:20:5) + 30 mg PEG-coated AuNPs mL^{-1} was observed. In contrast, the use of PNIPAM-coated AuNPs resulted in highly stable nanogels with no (<1%) AuNP release over time in vitro despite the high AuNP concentration in the gel (Figure S3, Supporting Information). Concentrations as high as 100 mg PNIPAM-coated AuNPs mL^{-1} were evaluated without induced burst release of encapsulated PNIPAM-coated AuNPs (data not shown). The enhanced compatibility of the PNIPAM-coated AuNPs with the SAIB matrix compared to the PEG-coated counterpart can be explained by the enhanced hydrophobicity and potential hydrogen bonding interactions between

ADVANCED HEALTHCARE MATERIALS
www.advhealthmat.de
www.MaterialsViews.com

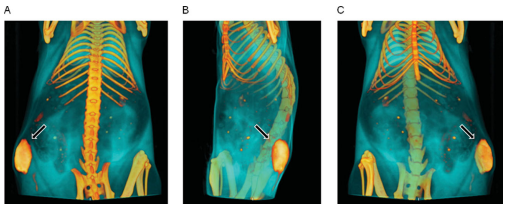


Figure 3. Representative 3D reconstruction from micro-CT images of mouse. Nanogel composed of SAIB/EtOH/PLA (75:20:5) + 30 mg PNIPAM-coated AuNPs mL^{-1} administered subcutaneously to immunocompetent mice at the upper left flank. Nanogel visualized from different angles.

the PNIPAM-coating polymers and the polyester backbone of SAIB. The positive in vitro stability results and known target AuNP concentration for 2D X-ray visibility encouraged us to further evaluate the nanogels composed of SAIB/EtOH/PLA (75:20:5) + 30 mg PNIPAM-coated AuNPs mL^{-1} in vivo using immunocompetent mice.

2.4. In Vivo Evaluation of Nanogel Performance

2.4.1. Evaluation of Nanogel in Immunocompetent Healthy NMRI-Mice

Nanogels (200 µL) composed of SAIB/EtOH/PLA (75:20:5) + 30 mg PNIPAM-coated AuNPs mL^{-1} were injected subcutaneously in immunocompetent NMRI-mice ($n = 5$) at the upper left flank using a 25G needle under general anesthesia. The nanogel was visualized using micro-CT imaging over a period of 8 weeks (images acquired directly after injection and again after 24, 48 h, 1, 2, 4, and 8 weeks) using a dedicated small-animal scanner (**Figure 3**). Based on the recorded micro-CT images, the nanogel X-ray contrast level, degradation rate, and homogeneity were evaluated as a function of time (**Figure 4**). No significant changes in the nanogel volume were observed after the initial EtOH efflux phase, which caused approximately 20% reduction of nanogel volume in good correlation with the known EtOH content of the nanogel prior to injection. The nanogel doped with 30 mg PNIPAM-coated AuNPs mL^{-1} provided a mean X-ray contrast level in the range of 1200 HU, which is sufficient for 2D X-ray visualization as previously described. A reduction of the mean contrast level of the nanogel was observed over time possible due to some AuNP leakage or aggregation of particles within the nanogel. Figure 4B clearly displays the increased maximum contrast level within the nanogel during the first week following intra-gel AuNP aggregation resulting in very high contrast levels (>15.000 HU) (Figure S4, Supporting Information). The homogeneity of the nanogel in terms of PNIPAM-coated AuNP distribution within the gel matrix was analyzed further by investigating each voxel of the micro-CT images as illustrated in Figure 4C. As evident from the cumulative histograms for a representative mouse, the PNIPAM-coated AuNPs were found to gradually aggregate within the nanogel as previously observed for PEG-coated AuNPs within the same gel matrix.[17] In fact, no significant difference was found in the variance of the PEG-coated and PNIPAM-coated AuNPs within the nanogels as analyzed by the Wilcoxon rank-sum test ($p = 0.0734$).[31] The inhomogeneous PNIPAM-coated AuNPs distribution within the nanogel could be visualized in the mice due to the high resolution and small slice thickness ($0.092 \times 0.092 \times 0.092$ mm^3) of the micro-CT scanner. The aggregation phenomenon of the PNIPAM-coated AuNPs within the nanogel has no clinical impact as human clinical scanners operate with significant larger slice thickness (2–3 mm) and lower resolution (1×1 mm^2) causing the AuNP microstructure aggregation to be below the resolution of recorded images.

2.4.2. Evaluation of Nanogel in a Canine Cancer Patient with Spontaneously Developed Solid Tumor

In order to fully validate the concept of liquid injectable fiducial markers for IGRT, the nanogel was evaluated in a canine cancer patient suffering from spontaneously developed malignant solid tumor (mast cell tumor). Solid tumors often show an increased interstitial fluid pressure (IFP) possibly due to vessel abnormalities, fibrosis, and contraction of the interstitial matrix mediated by stromal fibroblasts.[32,33] Increased IFP may serve as an obstacle for intratumoral injection of the developed nanogel as the available intracellular space for marker assembly post-injection is limited and the increased IFP may repel the nanogel into adjacent areas. Additionally, migration of the nanogel over time may compromise the accuracy of fiducial

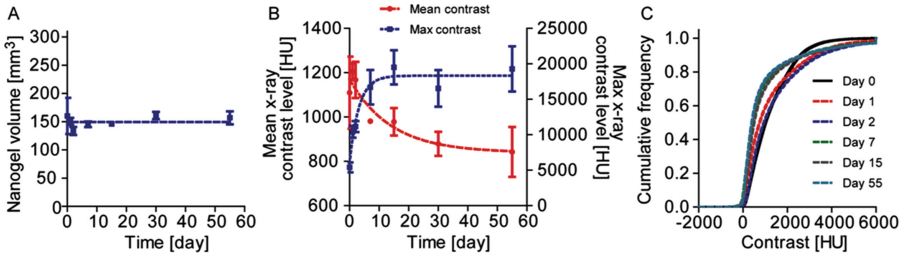


Figure 4. In vivo characteristics of nanogel composed of SAIB/EtOH/PLA (75:20:5) + 30 mg PNIPAM-coated AuNPs mL^{-1} administered subcutaneously to immunocompetent mice at the upper left flank. A) Nanogel degradation over a period of 8 weeks (mean ± SEM, $n = 5$); B) Mean (left axis) and maximum (right axis) X-ray contrast level of nanogel over a period of 8 weeks (mean ± SEM, $n = 5$) and C) Accumulated histogram of nanogel contrast level within each pixel from a representative mouse over a period of 30 d.

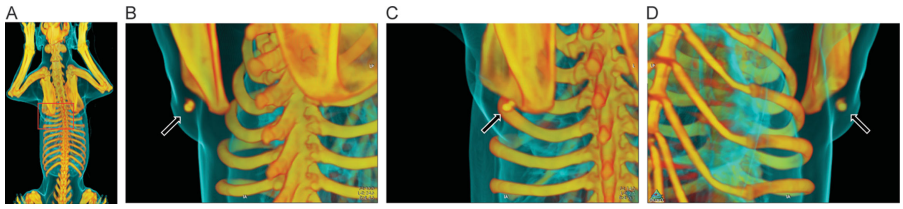


Figure 5. 3D reconstruction based on CT images of canine patient with injected nanogel. Nanogel composed of SAIB/EtOH/PLA (75:20:5) + 30 mg PNIPAM-coated AuNPs mL^{-1} administered intratumorally into a canine suffering from an intermediate-grade subcutaneous mast cell tumor (maximum distance ($x \times y \times z$); $1.82 \times 5.81 \times 5.32$ cm^3, CT volume; 31.64 cm^3) adherent to the underlying soft tissue located over the dorsal aspect of the left shoulder blade. A) Full-body scan of the canine. Area of interest indicated with a red box. B–D) Nanogel visualized from different angles.

marker-based IGRT and potentially result in misguidance of the external radiation beam, which, in turn, leads to poor tumor coverage and increased radiation to healthy adjacent tissue.[16,34] Such obstacles are not encountered for subcutaneous injections of the nanogel as evaluated in immunocompetent NMRI-mice, hence the canine cancer patient with a spontaneously developed solid tumor (mast cell tumor) served as a clinical animal model with high translational value for evaluating the clinical potential of the developed nanogel.

A companion dog (Male, Rhodesian Ridgeback, 10 years, body weight 38 kg) diagnosed with a subcutaneous mast cell tumor (maximum distance ($x \times y \times z$); $1.82 \times 5.81 \times 5.32$ cm^3, CT volume; 31.64 cm^3) adherent to the underlying soft tissue located over the dorsal aspect of the left shoulder blade was given an intratumoral injection of nanogel (300 µL) composed of SAIB/EtOH/PLA (75:20:5) + 30 mg PNIPAM-coated AuNPs mL^{-1} under general anesthesia (**Figure 5**) prior to radiation therapy. Approximately 24 h after the injection of the nanogel, a diagnostic planning CT-scan of the tumor/nanogel region was performed and the therapeutic regimen consisting of four fractions of 6 Gy radiation therapy planned accordingly. Radiation therapy was delivered over 16 d on a clinical linear accelerator equipped with on-board imaging (OBI) capabilities (3D cone beam CT (CBCT) and 2D X-ray imaging).

Prior to each fraction of radiotherapy, CBCT (**Figure 6**) images and 2D X-ray imaging (Figure S5, Supporting Information) were acquired and used for positioning based on nanogel and tumor. Radiation therapy was delivered according to the nanogel position. The size (**Table 1**) and stability of the nanogel in terms of migration and deformation were evaluated based on the recorded CBCT images (Figure 6) by manually contouring the nanogel using 300 HU as a lower threshold. The nanogel size remained constant over the time of treatment with no change in overall volume was observed. The small variations observed in Table 1 are a consequence of motion artifacts from respiratory movement on the recorded CBCT images. Additionally, image analysis revealed that the nanogel retained its 3D shape and was found to be immobilized/associated with the tumor as no migration of the nanogel was observed during the course of radiation therapy despite the canine patient exercising and resting on the tumor inducing mechanical stress on the nanogel. Finally, exposure to four fractions of 6 Gy each was not found to alter the nanogel morphology as evaluated based on the CBCT images. The nanogel was well tolerated and no clinical side effects were observed throughout the study.

3. Conclusion

A novel nanogel composition based on PNIPAM-coated AuNPs formulated in a mixture of SAIB/EtOH/PLA (70:20:5) has been developed. PNIPAM-coated AuNPs were found to be highly compatible with the hydrophobic SAIB-based gel matrix facilitating encapsulation of high concentration of colloidal gold, which enables 2D X-ray visualization of the nanogel. This was confirmed by a simulation of 2D DRRs of a patient suffering from lung cancer. The developed nanogel is the first injectable fiducial marker that can be visualized in 2D X-ray imaging thereby enabling intra- and interfractional IGRT by referencing the nanogel position. The nanogel performance was evaluated

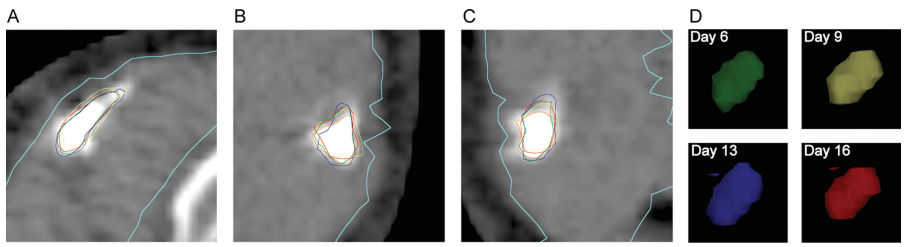


Figure 6. Nanogel composed of SAIB/EtOH/PLA (75:20:5) + 30 mg PNIPAM-coated AuNPs mL^{-1} administered intratumorally into a canine mast cell tumor visualized by CBCT imaging. A) Transversal plane (day 0); B) frontal plane (day 0); C) sagittal plane (day 0). Contours in A–C represent the nanogel on day 6 (green), day 9 (yellow), day 13 (blue), and day 16 (red). The light blue contour represents the gross tumor volume. D) 3D renderings of all the four surfaces represent the nanogel from each day of treatment.

ADVANCED HEALTHCARE MATERIALS
www.advhealthmat.de
www.MaterialsViews.com

Table 1. Nanogel volume following intratumoral injection in a canine model.

Day #	Nanogel volume [cm^3]
0	0.20
6	0.18
9	0.21
13	0.22
16	0.19

in vivo in immunocompetent mice by subcutaneous injection. Excellent visibility of the nanogel by micro-CT imaging was observed with a mean contrast level of 1200 HU after injection. Over time, the PNIPAM-coated AuNPs were found to form microaggregates within the gel matrix as confirmed by image analysis of the high-resolution micro-CT images. This phenomenon was found to have no clinical impact due to the larger slice thickness of human clinical scanners and lower resolution compared to small animal micro-CT scanners.

The concept of an injectable fiducial marker for use in IGRT was finally validated in a canine cancer patient with a large spontaneous solid tumor. The canine patient provides translational understanding and allowed us to evaluate the clinical potential of the nanogel. The nanogel displayed high image contrast on both conventional CT and CBCT images and could additionally be identified by 2D X-ray imaging, thus providing sufficient radiographic properties for visualization with all the imaging modalities associated with modern linear accelerators. Furthermore, the nanogel retained its size and 3D structure within the tumor microenvironment and no migration of the nanogel was observed within the time period for the radiotherapy treatment procedure. The developed nanogel was utilized to align the canine patient prior to external beam radiation therapy, was found to be unaffected by the deposited radiation dose within the tumor, and no side effects were observed in neither the murine nor the clinically relevant canine model. However, it should be noted that AuNPs are currently not approved for human use and the knowledge about the potential short- and long-term toxicological effects after AuNP exposure is still limited.[36]

The developed nanogel presented in this article is the first liquid fiducial marker, which can be inserted using thin and flexible needles and with a high enough contrast to make it visible in 2D X-ray imaging, which is an important feature for use in IGRT. Furthermore, active pharmaceutical ingredients may be included in the nanogel formulation to improve the therapeutic benefit for the patient, and the AuNPs may enhance the effect of the radiation therapy if the gel is injected directly into tumors.[37,38] In conclusion, the nanogel may provide both human and veterinary radiation oncologists with a new tool for improving planning, alignment, and tumor tracking in radiotherapy and could improve treatment outcome by providing a more precise delivery of the radiation dose to tumors.

4. Experimental Section

Materials: Tetrachloroauric acid trihydrate was purchased from Wako Chemicals GmbH (Neuss, Germany) and MeO-PEG_{5000}-SH was purchased from Rapp Polymere (Tuebingen, Germany). SH-PNIPAM (M_w = 3500Da, PDI = 1.24) was purchased from Polymer Source Inc. (Dorval, Canada). All other chemicals including SAIB, PLA (M_w = 10–18 kDa), trisodium citrate dihydrate, Au-standard for inductively coupled plasma atomic emission spectroscopy (ICP-AES) (1000 mg L^{-1}), hydrochloride acid and nitric acid were purchased from Sigma–Aldrich (Schnelldorf, Germany). All chemicals were used as received from the manufacturer.

*Synthesis of $(SH)_2$-SAIB (**5**)*: $(SH)_2$-SAIB (**5**) was synthesized over four steps starting from sucrose (**1**) in an overall yield of 27%. Full synthetic procedure including chemical characterization (^{1}H NMR, ^{13}C NMR, HRMS, and FTIR) is available in the Supporting Information.

AuNP Synthesis and Characterization: AuNPs were synthesized as previously described using a two-step seeding procedure resulting in citrate-stabilized AuNPs with an average diameter of 40 nm.[17] The citrate-stabilized AuNPs were subsequently coated with either SH-PEG_{5000}, SH-$PNIPAM_{3500}$, or $(SH)_2$-SAIB (**5**) using eight molecules per nm^2 surface area. Chemisorption of coating materials was in all cases conducted overnight at RT after which the volume was reduced to approximately 15 mL by centrifugation (9.500 RPM, 15 min). The PEG-coated AuNPs, PNIPAM-coated AuNPs, and SAIB-coated AuNPs colloid solutions were washed with MQ-H_2O (3 × 15 mL) and the finally up-concentrated to approximately 75 mg Au mL^{-1} as determined by ICP-MS using an ICP-MS Au standard (1000 mg L^{-1}) as reference. Colloid solutions (citrate-stabilized AuNPs, PEG-, PNIPAM-, and SAIB-coated AuNPs) were furthermore characterized by UV–vis spectrometry, DLS, and ζ-potential measured. PNIPAM-coated AuNPs were finally characterized by TEM (Figure S2, Supporting Information).

Nanogel Formulation: SAIB (1.75 g, 90 w/w% in EtOH) and PLA (105 mg) were mixed and homogenized by a ball homogenizer (frequency 30 s^{-1}, 30 min) to give a clear homogenous solution. Lyophilized PNIPAM-coated AuNPs (63 mg) was weighed off and redispersed into anhydrous EtOH (310 µL) by extensive vortexing and added to the SAIB/PLA mixture to give a final nanogel forming matrix composed of SAIB/EtOH/PLA (75:20:5) with 3.0 w/w% (30 mg mL^{-1}) PNIPAM-coated AuNPs. Formed nanogels were stored at RT and used within 4 h from preparation with vortexing prior to use. Nanogels with PEG-coated AuNPs were prepared as described elsewhere.[17]

In Vitro Release of PNIPAM- and PEG-Coated AuNPs from Nanogel: Nanogels composed of SAIB/EtOH/PLA (75:20:5) with 3.0 w/w% (30.0 mg mL^{-1}) of either PEG- or PNIPAM-coated AuNPs were prepared as described above. Nanogels (200 µL) were added to glass vials containing MQ-H_2O (10.0 mL, 37 °C) and aliquots (1.0 mL) were removed as a function of time and replaced with an equal amount of MQ-H_2O. The release profile from both nanogels was measured by correlating the UV–vis absorbance of the individual AuNPs with a standard curve based on Au-standards of known concentration (5.0–50.0 µg mL^{-1}).

Simulation of CNR in Lung Cancer Patient: 2D DRRs were performed based on clinical 3D CT-data for a patient diagnosed with lung cancer. 2D DRRs were created using the drrCode ray-tracing algorithm available online (https://code.google.com/p/drrsuite/downloads/list). 2D DRRs were created using a source to detector distance of 150 cm and a detector size of 768 × 1024 pixels (identical to a clinical settings). For conversion from HU to attenuation coefficient, a μ_ρ for water was used assuming monochromatic photons of 120 kV. ROI contouring and CNR calculations for the simulated projections were performed in MATLAB (The MathWorks Inc., Natick, MA, USA) using an in-house-developed simulation and analysis tool.

Animal Models (Mice): Immunocompetent 7-week-old female NMRI mice (n = 5) were given a subcutaneous injection (200 µL) of SAIB/EtOH/PLA (75:20:5) + 30 mg mL^{-1} PNIPAM-coated AuNPs using a 25 G needle in the upper left flank of the anesthetized mice. The animals were given free access to food and water and micro-CT scans were recorded over a period of 8 weeks (24, 48 h, 1, 2, 4, and 8 weeks) using a dedicated small-animal MicroCAT II system (Siemens Medical Solutions, Malvern, PA, USA). Micro-CT images were recorded using the following settings; tube voltage of 67 kVp, tube current of 500 µA, 360 rotation steps, an exposure time of 400 ms, and a voxel size of 0.092 × 0.092 × 0.092 mm^3. The animal experiments were approved by the Animal Research Committee of the Danish Ministry of Justice.

Canine: A privately owned companion dog (Male, Rhodesian Ridgeback, 10 years, bodyweight 38 kg) suffering from a spontaneously developed subcutaneous mast cell tumor (maximum distance ($x \times y \times z$); $1.82 \times 5.81 \times 5.32$ cm^3, CT volume; 31.64 cm^3) adherent to the underlying soft tissue located over the dorsal aspect of the left shoulder blade was used as a clinical relevant model. In conjunction with the standard diagnostic procedures, nanogel (300 µL) was manually injected centrally in the tumor using a 23 G needle and 1 mL syringe. Approximately 24 h after the injection of the nanogel, a diagnostic CT scan (Biograph 40, Siemens Medical Solutions, Erlangen, Germany) of the tumor region was performed with the dog under general anesthesia using the following acquisition parameters; 120 kV, 150 mAs, pixel size $0.977 \times 0.977 \times 1.00$ mm^3). As part of the therapeutic procedure radiation therapy was planned according to the tumor histology and imaging characteristic, and a therapeutic regimen consisting of four fractions of 6 Gy radiation treatments was chosen. Radiation therapy was delivered by a linear accelerator (Novalis TX, Varian Medical Systems, Palo Alto, CA, USA) with on-board imaging capabilities (Cone Beam CT and 2D X-ray imaging) and orthogonal ExacTrac radiography (Brainlab AG, Feldkirchen, Germany). No side effects were observed throughout the study. The canine cancer patient study was approved by the Ethics and Administrative Committee at the Department of Veterinary Clinical and Animal Sciences, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark.

Analysis of Nanogel Homogeneity: The nanogels were delineated by hand using ITK-SNAP.[35] All data analyses were done in MATLAB (The MathWorks Inc., Natick, MA, USA) using in-house- developed software. Based on the voxel values in the gels mean and max values, as well as cumulative histograms were calculated. The cumulative histograms were used for homogeneity evaluation, as changes in the AuNP distribution are reflected in the voxels values and thus the histograms.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

This work has been funded by the Technical University of Denmark (DTU), the Niels Bohr Institute, the Danish Council for Strategic Research (Nanoguide, application no. 0603-00442B), and the Novo Foundation.

Received: October 17, 2014
Revised: December 3, 2014
Published online: January 21, 2015

[1] R. Baskar, K. A. Lee, R. Yeo, K. Yeoh, *Int. J. Med. Sci.* **2012**, *9*, 193.
[2] L. A. Dawson, M. B. Sharpe, *Lancet Oncol.* **2006**, *7*, 848.
[3] J. Bussink, J. H. A. M. Kaanders, W. T. A. van der Graaf, W. J. G. Oyen, *Nat. Rev. Clin. Oncol.* **2011**, *8*, 233.
[4] D. Verellen, R. M. De, N. Linthout, K. Tournel, G. Soete, G. Storme, *Nat. Rev. Cancer* **2007**, *7*, 949.
[5] K. M. Langen, D. T. Jones, *Int. J. Radiat. Oncol. Biol. Phys.* **2001**, *50*, 265.
[6] H. Shirato, Y. Seppenwoolde, K. Kitamura, R. Onimura, S. Shimizu, *Semin. Radiat. Oncol.* **2004**, *14*, 10.
[7] P. Giraud, E. Yorke, S. Jiang, L. Simon, K. Rosenzweig, G. Mageras, *Cancer Radiother.* **2006**, *10*, 269.
[8] M. Falk, P. M. af Rosenschold, P. Keall, H. Cattell, B. C. Cho, P. Poulsen, S. Povzner, A. Sawant, J. Zimmerman, S. Korreman, *Radiother. Oncol.* **2010**, *94*, 218.
[9] D. P. Harley, W. S. Krimsky, S. Sarkar, D. Highfield, C. Aygun, B. Gurses, *Ann. Thorac. Surg.* **2010**, *89*, 368.
[10] M. Imura, K. Yamazaki, H. Shirato, R. Onimaru, M. Fujino, S. Shimizu, T. Harada, S. Ogura, H. Dosaka-Akita, K. Miyasaka, M. Nishimura, *Int. J. Radiat. Oncol. Biol. Phys.* **2005**, *63*, 1442.
[11] D. J. Moseley, E. A. White, K. L. Wiltshire, T. Rosewall, M. B. Sharpe, J. H. Siewerdsen, J. Bissonnette, M. Gospodarowicz, P. Warde, C. N. Catton, D. A. Jaffray, *Int. J. Radiat. Oncol. Biol. Phys.* **2007**, *67*, 942.
[12] I. Gauthier, J. Carrier, D. Beliveau-Nadeau, B. Fortin, D. Taussky, *Int. J. Radiat. Oncol. Biol. Phys.* **2009**, *74*, 1128.
[13] R. Khosa, S. Nangia, K. S. Chufal, D. Ghosh, R. Kaul, L. Sharma, *J. Cancer Res. Ther.* **2010**, *6*, 172.
[14] S. Yousefi, B. T. Collins, C. A. Reichner, E. D. Anderson, C. Jamis-Dow, G. Gagnon, S. Malik, B. Marshall, T. Chang, F. Banovac, *Clin. Lung Cancer* **2007**, *8*, 252.
[15] N. Kothary, J. J. Heit, J. D. Louie, W. T. Kuo, B. W. J. Loo, A. Koong, D. T. Chang, D. Hovsepian, D. Y. Sze, L. V. Hofmann, *J. Vasc. Interv. Radiol.* **2009**, *20*, 235.
[16] N. Bhagat, N. Fidelman, J. C. Durack, J. Collins, R. L. Gordon, J. M. LaBerge, R. K. J. Kerlan, *Cardiovasc. Interv. Radiol.* **2010**, *33*, 1186.
[17] R. I. Jølck, T. Binderup, A. E. Hansen, J. B. Scherman, P. M. af Rosenschold, A. Kjær, T. L. Andresen, *Adv. Healthcare Mater.* **2014**, *3*, 1680.
[18] H. Shirato, K. Suzuki, G. C. Sharp, K. Fujita, R. Onimaru, M. Fujino, N. Kato, Y. Osaka, R. Kinoshita, H. Taguchi, S. Onodera, K. Miyasaka, *Int. J. Radiat. Oncol. Biol. Phys.* **2006**, *64*, 1229.
[19] T. Juhler-Nottrup, S. S. Korreman, A. N. Pedersen, G. F. Persson, L. R. Aarup, H. Nystrom, M. Olsen, N. Tarnavski, L. Specht, *Acta Oncol.* **2008**, *47*, 1406.
[20] S. S. Korreman, T. Juhler-Nottrup, G. F. Persson, P. A. Navrsted, M. Enmark, H. Nystrom, L. Specht, *Acta Oncol.* **2008**, *47*, 1390.
[21] M. L. Taylor, T. Kron, R. D. Franich, *Acta Oncol.* **2011**, *50*, 483.
[22] C. Rubio, O. Hernando, R. Morera, T. Leroy, S. E. Lartigau, *Rep. Pract. Oncol. Radiother.* **2013**, *18*, 387.
[23] Y. Lu, H. He, Y. Cui, X. Tang, *Yao Xue Xue Bao.* **2007**, *42*, 445.
[24] Y. Lu, Y. Yu, X. Tang, *J. Pharm. Sci.* **2007**, *96*, 3252.
[25] X. Lin, S. Yang, J. Gou, M. Zhao, Y. Zhang, N. Qi, H. He, C. Cai, X. Tang, P. Guo, *J. Mater. Sci.: Mater. Med.* **2012**, *23*, 443.
[26] M. M. Andrade, M. T. Barros, *Tetrahedron* **2004**, *60*, 9235.
[27] M. Brust, M. Walker, D. Bethell, D. J. Schiffrin, R. Whyman, *Chem. Commun.* **1994**, *7*, 801.
[28] F. Jacobs, E. Sundermann, B. De Sutter, M. Christiaens, I. Lemahieu, *J .Comput. Inf. Technol.* **1998**, *6*, 89.
[29] R. L. Siddon, *Phys. Med. Biol.* **1985**, *30*, 817.
[30] Y. Chen, J. J. O'Connell, C. J. Ko, R. R. Mayer, A. Belard, J. E. McDonough, *Phys. Med. Biol.* **2012**, *57*, 155.
[31] F. Wilcoxon, *J. Econ. Entomol.* **1946**, *39*, 269.
[32] C. Heldin, K. Rubin, K. Pietras, A. Oestman, *Nat. Rev. Cancer* **2004**, *4*, 806.
[33] P. Vaupel, *Semin. Radiat. Oncol.* **2004**, *14*, 198.
[34] J. C. Hong, N. C. W. Eclov, Y. Yu, A. K. Rao, S. Dieterich, Q. Le, M. Diehn, D. Y. Sze, B. W. Loo Jr., N. Kothary, P. G. Maxim, *J. Appl. Clin. Med. Phys.* **2013**, *14*, 4046.
[35] P. A. Yushkevich, J. Piven, H. C. Hazlett, R. G. Smith, S. Ho, J. C. Gee, G. Gerig, *Neuroimage* **2006**, *31*, 1116.
[36] A. Gerber, M. Bundschuh, D. Klingelhofer, D. A. Groneberg, *J. Occup. Med. Toxicol.* **2013**, *8*, 32.
[37] X. Zhang, D. Wu, X. Shen, J. Chen, Y. Sun, P. Liu, X. Liang, *Biomaterials* **2012**, *33*, 6408.
[38] X. Zhang, J. Chen, Z. Luo, D. Wu, X. Shen, S. Song, Y. Sun, P. Liu, J. Zhao, S. Huo, S. Fan, F. Fan, X. Liang, J. Xie, *Adv. Healthcare Mater.* **2014**, *3*, 133.

ADVANCED HEALTHCARE MATERIALS

Supporting Information

for *Adv. Healthcare Mater.*, DOI: 10.1002/adhm.201400651

Injectable Colloidal Gold for Use in Intrafractional 2D Image-Guided Radiation Therapy

*Rasmus I. Jølck, Jonas S. Rydhög, Anders N. Christensen, Anders E. Hansen, Linda M. Bruun, Henrik Schaarup-Jensen, Asger Stevner von Wenck, Betina Børresen, Annemarie T. Kristensen, Mads H. Clausen, Andreas Kjær, Knut Conradsen, Rasmus Larsen, Per Munck af Rosenschöld, and Thomas L. Andresen**

WILEY-VCH

Supporting Information

Injectable Colloidal Gold for use in Intrafractional 2D Image Guided Radiation Therapy

*Rasmus I. Jølck, Jonas S. Rydhög, Anders N. Christensen, Anders E. Hansen, Linda M. Bruun, Henrik Schaarup-Jensen, Asger Stevner von Wenck, Betina Børresen, Annemarie T. Kristensen, Mads H. Clausen, Andreas Kjær, Knut Conradsen, Rasmus Larsen, Per Munck af Rosenschöld and Thomas L. Andresen**

Dr. R. I. Jølck, Dr. A. E. Hansen, L. M. Bruun and Prof. T. L. Andresen.
DTU Nanotech, Department of Micro-and Nanotechnology, Center for Nanomedicine and Theranostics, Technical University of Denmark, Building 345E, Ørsteds Plads, 2800 Lyngby, Denmark.
E-mail: thomas.andresen@nanotech.dtu.dk

J. S. Rydhög, Dr. P. Munck af Rosenschöld
Department of Oncology, Section of Radiotherapy, 3994, Rigshospitalet, Blegdamsvej 9, 2100 Copenhagen, Denmark and the Niels Bohr Institute, University of Copenhagen, Blegdamsvej 17, 2100 Copenhagen, Denmark.

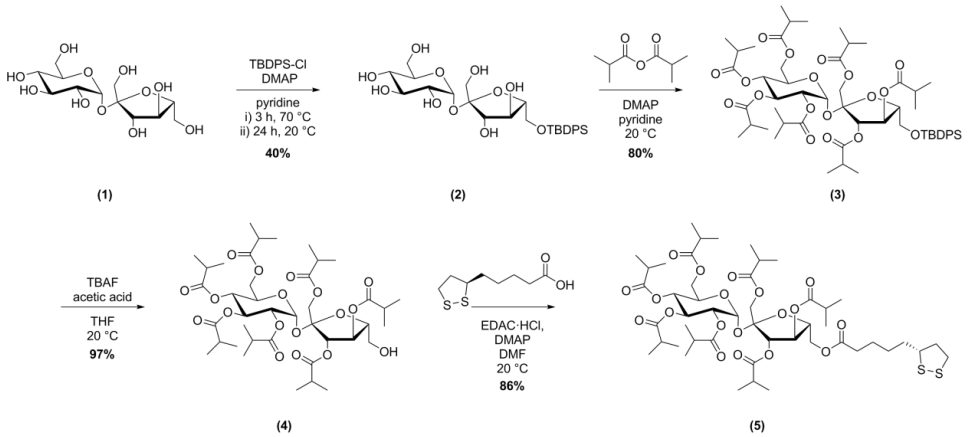
A. N. Christensen, Prof. K. Conradsen and Prof. R. Larsen.
DTU Compute, Department of Applied Mathematics and Computer Science, Technical University of Denmark, Building 321/324, Matematiktorvet, 2800 Lyngby, Denmark.

H. Schaarup-Jensen and Prof. M. H. Clausen.
DTU Chemistry, Department of Chemistry, Center for Nanomedicine and Theranostics, Technical University of Denmark, Building 207, Kemitorvet, 2800 Lyngby, Denmark.

Prof. A. Kjær
Department of Clinical Physiology, Nuclear Medicine & PET and Cluster for Molecular Imaging, Rigshospitalet and University of Copenhagen, Blegdamsvej 9, 2100 Copenhagen, Denmark.

Prof. A. T. Kristensen, A. Stevner von Wenck and B. Børresen
Department of Veterinary Clinical and Animal Sciences, Faculty of Health and Medical Sciences, University of Copenhagen, Dyrlægevej 16, 1870 Frederiksberg C, Denmark.

1. Synthetic Procedure to $(SH)_2$-SAIB and Product Characterization



Scheme S1. Synthesis scheme for obtaining $(SH)_2$-SAIB (**5**) over four synthetic steps starting from sucrose (**1**) resulting in $(SH)_2$-SAIB (**5**) in an overall yield of 27%.

General Information

Nuclear Magnetic Resonance (NMR) was conducted on a Bruker Ascend^{TM} 400 MHz - operating at 401.3 MHz for ^{1}H and 100.62 MHz for ^{13}C - with a 5 mm H – Broadband Dual Channel z-gradient Prodigy cryoprobe instrumentation. All spectrums were measured at 298 K and the spectrums recorded were calibrated according to the solvent residues. Thin Layer Chromatography (TLC) analysis was conducted by use of TLC Silica gel 60 F_{254} aluminum sheets, Merck KGaA, 64271 Darmstadt, Germany. As developing agent a mixture (Cemol-dip) consisting of 9 g cerium sulfate, 47 g ammonium molybdate, 849 mL H_2O and 94 mL of sulfuric acid was applied on the TLC plate followed by heat treatment. For dry column purification, a silica gel 60 (0.015-0.040 mm) from Merck KGaA, 64271 Darmstadt, Germany was used. High Resolution Mass spectroscopy (HR-MS) was conducted on a Bruker Daltronics maXis 3G QTOF-MS fitted with a Dionex Ultimate 3000 UHPLC instrumentation. Attenuated Total Reflection Fourier Transformation Inferred radiation (ATR-FTIR) analysis was conducted on a Bruker Alpha platinum ATR instrumentation with a single reflection diamond ATR module. Melting point was measured with a Stuart SMP 30 instrumentation. Sucrose (99.5%, 57-50-1), 4-(dimethylamino)pyridine (DMAP, puriss, 1122-58-3), isobutyric anhydride (97%, 97-72-3), (*R*)-(+)-1,2-dithiolane-pentanoic acid (97%, 1200-22-2), *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride (EDAC·HCl, 98%, 25952-53-8) and tetra-*n*-butylammoniumfluoride (TBAF, 1.0 M, 429-41-4) in THF were purchased from Sigma Aldrich (Steinheim, Germany) and used as received. *tert*-Butyl(chloro)diphenylsilane (TBDPS-Cl, 98%, 58479-61-1) was purchased from ABCR (Karlsruhe, Germany).

(2R,3R,4S,5S,6R)-2-(((2S,3S,4S,5R)-5-(((tert-Butyldiphenylsilyl)oxy)methyl)-3,4-dihydroxy-2-(hydroxymethyl)tetrahydrofuran-2-yl)oxy)-6-(hydroxymethyl)tetrahydro-2H-pyran-3,4,5-triol **(2)**. To a suspension of sucrose (**1**) (10 g, 29.2 mmol) in pyridine (100 mL) was added DMAP (0.36 g, 2.92 mmol). The reaction mixture was stirred at room temperature for 1 h followed by addition of TBDPS-Cl (4.65 mL, 17.5 mmol) and heated to 70 ^{0}C for 3 h. Another portion of TBDPS-Cl (4.65 mL, 17.5 mmol) was added to the reaction mixture at room temperature and then left stirred overnight. The reaction was followed by TLC. Pyridine was then removed *in vacuo* and the crude compound was purified by dry column vacuum chromatography (EtOH in EtOAc with 2% increments). Yield: 6.8 g amorphous solid; 40%.

Mp: 174 °C. ^{1}H-NMR (400 MHz, MeOD): δ = 1.04 (s, 9H), 3.25-3.35 (m, 2H), 3.50-3.68 (m, 5H), 3.72-3.79 (m, 1H), 3.84-3.99 (m, 3H), 4.01-4.13 (m, 2H), 5.44 (d, *J*=3.9 Hz, 1H), 7.35-7.46 (m, 6H), 7.67-7.77 (m, 4H). ^{13}C-NMR (100 MHz, MeOD): δ = 20.0, 27.3 (3C), 62.4, 64.3, 66.5, 71.4, 73.2, 73.9, 74.9, 76.2, 79.0, 83.7, 93.0, 105.6, 128.8 (2C), 128.8 (2C), 130.8 (2C), 134.5, 134.6, 136.8 (2C), 136.8 (2C); HR-MS. Calculated adducts: [M+H$^+$] = 581.2413, [M+Na$^+$] = 603.2232, [M+K$^+$] = 619.1971. Observed adducts, [M+H_2O] = 598.2675, [M+Na$^+$] = 603.2226, [M+K$^+$] = 619.1952. ATR-FTIR (cm^{-1}): 507.18, 701.83, 805.57, 930.50, 994.46, 1043.53, 1067.96, 1111.82, 1390.34, 1428.36, 1470.26, 2857.74, 2930.17, 3285.79.

(2R,3R,4S,5R,6R)-2-(((2S,3S,4R,5R)-5-(((tert-Butyldiphenylsilyl)oxy)methyl)-3,4-bis(iso-butyryloxy)-2-((isobutyryloxy)methyl)tetrahydrofuran-2-yl)oxy)-6-((isobutyryloxy)methyl-)tetrahydro-2H-pyran-3,4,5-triyl tris(2-methylpropanoate) **(3)**. 6'-O-TBDPS-sucrose (**2)** (2.00 g, 3.44 mmol) was dissolved into pyridine (40 mL) together with DMAP (0.114 g, 0.924 mmol). Isobutyric anhydride (11.4 mL, 66.7 mmol) was added and the reaction mixture was stirred at room temperature overnight. The reaction was followed by TLC. At the point of reaction completion, the reaction mixture was concentrated *in vacuo* and diethyl ether (200 mL) was added followed by extraction with saturated aqueous sodium bicarbonate solution (5×200 mL). The organic phase was subsequently collected and concentrated *in vacuo*. The crude product was purified by dry column purification (EtOAc in heptane with 2% increments). Yield = 2.97 g transparent oily texture; 80%.^{1}H-NMR (400 MHz, DMSO): δ = 0.94-1.07 (m, 39H), 1.08-1.16 (m, 12H), 2.32-2.48 (m, 4H), 2.52-2.62 (m, 3H), 3.75-3.86 (m, 2H), 3.90 (dd, *J* = 13 Hz, 2 Hz, 1H), 4.02-4.20 (m, 4H), 4.24 (ddd, *J* = 10.3 Hz, 4.4 Hz, 1.8 Hz, 1H), 4.86 (dd, *J* = 10.6 Hz, 3.6 Hz, 1H), 5.04 (t, *J* = 10.0 Hz, 1H), 5.35 (t, *J* = 10.0 Hz, 1H); 5.46-5.49 (m, 2H); 5.62 (d, *J* = 3.7 Hz, 1H), 7.37-7.50 (m, 6H), 7.56-7.64 (m, 4H).^{13}C-NMR (100 MHz, DMSO): δ = 18.0-19.0 (14C), 26.4 (3C), 30.7, 32.9-33.3 (7C), 61.2, 63.3, 63.8, 67.1, 68.1, 69.1, 69.2, 73.5, 75.3, 80.0, 89.1, 102.4, 127.9 (4C), 130.0 (2C), 132.3, 132.4, 135.0 (4C), 174.6, 175.0 (3C), 175.2, 175.3, 175.6,. HR-MS. Calculated adducts: [M+H$^+$] = 1071.5343, [M+Na$^+$] = 1093.5163, [M+K$^+$] = 1109.4902. Observed adducts, [M+H_2O] = 1088.5609, [M+Na$^+$] = 1093.5175, [M+K$^+$] = 1109.4892. ATR-FTIR (cm^{-1}): 504.26, 702.23, 742.56, 823.65, 1030.51, 1071.63, 1111.82, 1139.69, 1186.65, 1247.91, 1388.47, 1470.12, 1741.41, 2877.85, 2934.80, 2973.83, 3285.44.

(2R,3R,4S,5R,6R)-2-(((2S,3S,4R,5R)-5-(hydroxymethyl)-3,4-bis(isobutyryloxy)-2-(iso-butyryloxy)methyl)tetrahydrofuran-2-yl)oxy)-6-((isobutyryloxy)methyl)tetrahydro-2H-pyran-3,4,5-triyl tris(2-methylpropanoate) **(4)**.

Hepta-isobutyryl-6'-O-TBDPS-sucrose (**3)** (1.032 g, 0.95 mmol) was dissolved in dry THF (20 mL) under inert atmosphere. Acetic acid (0.082 mL, 1.43 mmol) was added to the stirred mixture followed by the addition of a 1.0 M TBAF solution in THF (1.2 mL, 1.2 mmol). The reaction mixture was stirred at room temperature overnight. The reaction was followed by TLC. The reaction mixture was directly purified by dry column purification (EtOAc in heptane with 4% increments). Yield: 0.768 g transparent oily texture; 97%. ^{1}H-NMR (400 MHz, DMSO): δ = 0.95-1.21 (m, 42 H), 2.34-2.66 (m, 7H (DMSO overlap)), 3.60 (m, 2H), 4.00-4.12 (m, 4H), 4.21 (dd, *J* = 13.0 Hz, 4.4 Hz, 1H), 4.34 (ddd, *J* = 10.4 Hz, 4.5 Hz, 1.8 Hz, 1H), 4.91 (dd, *J* = 10.4, 3.6 Hz, 1H), 4.96 (t, *J* = 5.5 Hz, 1H), 5.11 (t, *J* = 9.9 Hz, 1H), 5.34 (t, *J* = 7.5 Hz, 1H), 5.40 (t, *J* = 9.9 Hz, 1H), 5.47 (d, *J* = 7.5 Hz, 1H), 5.67 (d, *J* = 3.5 Hz, 1H). ^{13}C-NMR (100 MHz, DMSO): δ = 18.0-19.0 (14C), 32.9-33.3 (7C), 61.4, 61.5, 63.3, 67.3, 68.0, 69.2, 69.3, 73.7, 75.2, 80.8, 89.1, 102.2, 174.6, 175.0 (3C), 175.3, 175.4, 175.8. HR-MS. Calculated adducts: [M+H$^+$] = 833.4165, [M+Na$^+$] = 855.3985, [M+K$^+$] = 871.3724. Observed adducts: [M+H_2O] = 850.4449, [M+Na$^+$] = 855.3999, [M+K$^+$] = 871.3738. ATR-

FTIR (cm^{-1}): 748.36, 1005.66, 1048.73, 1077.10, 1142.42, 1186.51, 1247.10, 1344.51, 1388.51, 1470.09, 1739.30, 2878.48, 2939.02, 2974.80, 3531.46.

*(2R,3R,4S,5R,6R)-2-(((2S,3S,4R,5R)-5-(((5-((R)-1,2-dithiolan-3-yl)pentanoyl)oxy)methyl)-3,4-bis(isobutyryloxy)-2-((isobutyryloxy)methyl)tetrahydrofuran-2-yl)oxy)-6-((isobutyryloxy)-methyl)tetrahydro-2H-pyran-3,4,5-triyl tris(2-methylpropanoate) (***5**).
Hepta-isobutyryl-sucrose (**4)** (0.543 g, 0.64 mmol) was dissolved in dry DMF (10 mL) under inert atmosphere. In another flask under inert atmosphere, (*R*)-(+)-1,2-Dithiolane-pentanoic acid (0.206 g, 0.96 mmol), DMAP (0.246 g, 1.91 mmol) and EDAC·HCl (0.190 g, 0.96 mmol) were dissolved in dry DMF (10 mL) and stirred for 15 min. The mixture, containing the activated ester was subsequently transferred to the reaction mixture of hepta-isobutyryl-sucrose (**4).** The reaction mixture was left stirred at room temperature for reaction. The reaction was followed by TLC. After 24 h a new mixture of activated ester: (*R*)-(+)-1,2-Dithiolane-pentanoic acid (0.0706 g, 0.32 mmol), DMAP (0.080 g, 0.64 mmol) and EDAC·HCl (0.065 g, 0.32 mmol) dissolved in dry DMF (10 mL) under inert atmosphere was added to the reaction mixture. At the point of reaction completion, the reaction mixture was concentrated *in vacuo* followed by addition of diethyl ether (100 mL) and phase extraction with water (3×100 mL). The organic phase was collected and finally purified by dry column purification (EtOAc in heptane with 4% increment per fraction). Yield: 0.558 g yellowish oily texture; 86%. ^{1}H-NMR (400 MHz, DMSO): δ = 0.96-1.18 (m, 42H), 1.38 (m, 2H), 1.48-1.72 (m, 4H), 1.86 (dq, *J* = 13.4 Hz, 6.4 Hz, 1H), 2.33 (t, *J* = 7.3 Hz, 2H), 2.36-2.44 (m, 2H), 2.44-2.66 (m, 6H), 3.14 (m, 2H), 3.59 (dq, *J* = 8.5 Hz, 6.2 Hz, 1H), 4.03-4.14 (m, 3H), 4.16-4.40 (m, 5H), 4.91 (dd, *J* = 10.4 Hz, 3.6 Hz, 1H), 5.10 (t, *J* = 9.9 Hz, 1H), 5.34-5.43 (m, 2H), 5.52 (d, *J* = 7.5 Hz, 1H); 5.64 (d, *J* = 3.5 Hz, 1H). ^{13}C-NMR (100 MHz, DMSO): δ = 18.0-19.0 (14C), 24.1, 28.1, 33.0-33.3 (8C), 34.1, 38.1, 40.0 (DMSO overlap), 56.0, 61.4, 63.2, 63.3, 67.3, 68.2, 69.2, 69.2, 73.5, 74.8, 77.5, 89.3, 102.4, 172.4, 174.5, 174.9, 175.0 (2C), 175.2, 175.4, 175.8. HR-MS. Calculated adducts: $[M+H^+]$ = 1021.4495, $[M+Na^+]$ = 1043.4314, $[M+K^+]$ = 1059.4054. Observed adducts, $[M+H_2O]$ = 1039.4844, $[M+Na^+]$ = 1043.4355, $[M+K^+]$ = 1059.4136. ATR-FTIR (cm^{-1}): 476.19, 748.51, 849.06, 916.58, 963.11, 1028.24, 1074.25, 1140.82, 1185.60, 1246.17, 1344.85, 1388.11, 1469.77, 1738.20, 2877.68, 2936.72, 2974.16.

2. AuNP Characterization and Handling

General considerations

The synthesized AuNPs were diluted with MQ-H_2O (Au-seeds ×4), citrate stabilized AuNPs (×4) and PNIPAM-coated AuNPs (×4000) prior to UV-vis, DLS and ζ-potential measurements.

UV-vis Measurements

The UV-vis absorption spectra of the synthesized AuNPs were measured using a Unicam Helios α Uni 9423 UV-vis spectrophotometer by scanning the UV-range from 450-900 nm. UV-vis absorption spectra of hydrosols were measured using disposable PS cuvettes whereas organosols were measured in quarts cuvettes of identical dimensions (1.0 cm × 1.0 cm). The apparent solvent was used in the reference cell for the individual measurements and blanks were run prior to AuNP analysis.

Dynamic Light Scattering Measurements

The hydrodynamic diameter and polydispersity of the synthesized AuNPs were determined using a Zeta PALS Analyzer (Brookhaven Instruments, NY, USA). Data was fitted using built-in software to calculate the hydrodynamic size of the AuNP-colloids.

ζ-potential Measurements

The ζ-potential was measured in the same sample as the DLS analysis using a Zeta PALS Analyzer (Brookhaven Instruments, NY, USA) with a pre-conditioned electrode with 10 sub-runs while observing a fitting model residual less than 0.04.

ICP-MS Measurements

The Au-concentration of the purified and up-concentrated PEG- and PNIPAM-coated AuNPs stock solutions were determined by ICP-MS using a Thermo Scientific iCAPQ ICP-MS (MA, USA) using the following settings; RF Power: 1550 W, Plasma Gas Flow: 14L min^{-1}, Nebulizer Gas Flow: 0.945L min^{-1}, Dwell Time: 0.01 s, Sweeps: 300, Replicates: 3, Delay Time: 60 s, Wash Time: 60 s. PNIPAM-coated AuNPs (20.0 µL) was digested using fresh prepared aqua regia (400 µL) and the total volume adjusted to 6.00 mL with MQ-H_2O (5580 µL) (× 300 dilution). The ×300 dilution solution (5 µL) was further diluted with HNO_3 (4995 µL, 1% + 0.5 ppb Ir as internal standard) to give a final dilution factor of ×300.000. A standard curve covering the concentration range of 62.5 ppb ─ 500 ppb in HNO_3 (1% + 0.5 ppb Ir as internal standard) was prepared using an Au-standard for ICP-MS (1000 mg L^{-1}). The intensity of ^{197}Au (m/z = 196.9666) was measured and the concentration of coated AuNPs was determined based on the corresponding standard curve.

Lyophilization and Stability of PNIPAM-coated AuNPs

The PNIPAM-coated AuNPs stock solution (64 mg Au mL^{-1}) was diluted to 1.0- 2.5- and 5.0 mg Au mL^{-1} with MQ-H_2O in glass vials. The samples were snap frozen by submerging the samples into liquid nitrogen and lyophilized overnight ($p < 6.0 \times 10^{-2}$ mbar) to give an air stable black nanopowder which was stored in the dark at 5 °C until further use. The lyophilized PNIPAM-coated AuNPs readily re-dispersed in MQ-H_2O to form the apparent PNIPAM-coated AuNPs hydrosol with no change in particle morphology based on UV-vis spectroscopy and DLS analysis. Additionally, the lyophilized PNIPAM-coated AuNP nanopowder also readily re-dispersed in various protic- and aprotic organic solvents to form stable PNIPAM-coated AuNP organosols (**Figure S1**).

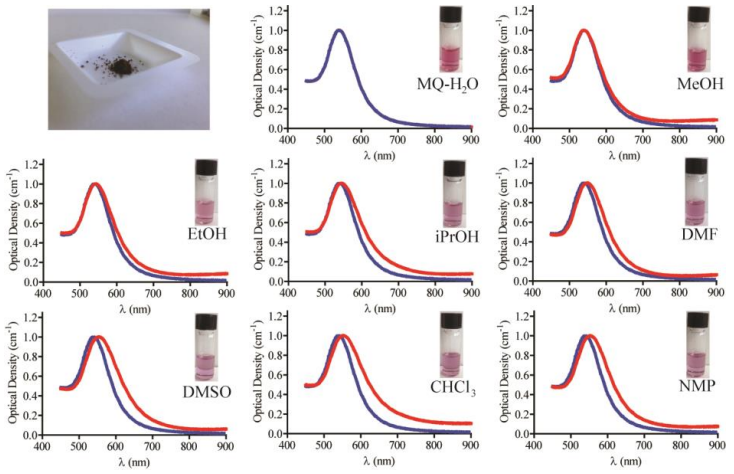


Figure S1. UV-vis spectroscopy analysis of the re-dispersibility of the air stable black nanopowder of PNIPAM-coated AuNPs obtained by lyophilization in MQ-H_2O and various protic- and aprotic polar organic solvents covering the relative polarity scale from 0.259 (chloroform) to 1.000 (MQ-H_2O). An apparent red-shift of the surface plasmon resonance peak is observed following re-dispersion in organic solvent due to the change in refractive index of the solvent.

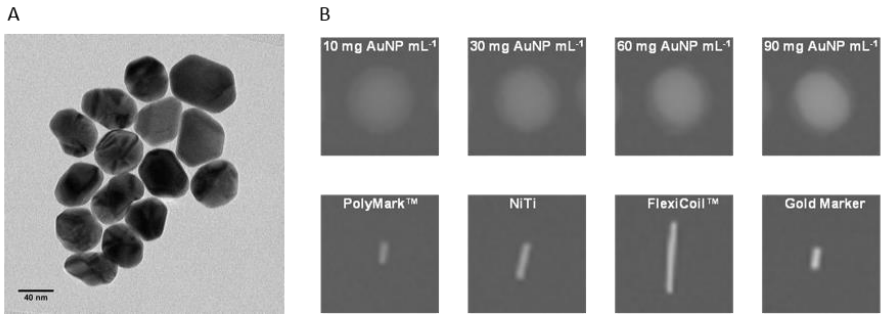


Figure S2. A) Characterization of PNIPAM-coated AuNP by TEM; B) Comparison of the obtained contrast level in 2D x-ray imaging of AuNPs containing nanogels (10- 30-, 60- and 90 mg AuNP mL^{-1}) and solid fiducial markers used in the clinic. Markers visualized in a 10 cm solid water phantom using the following settings; 80 kVp, 80 mA, 80 ms.

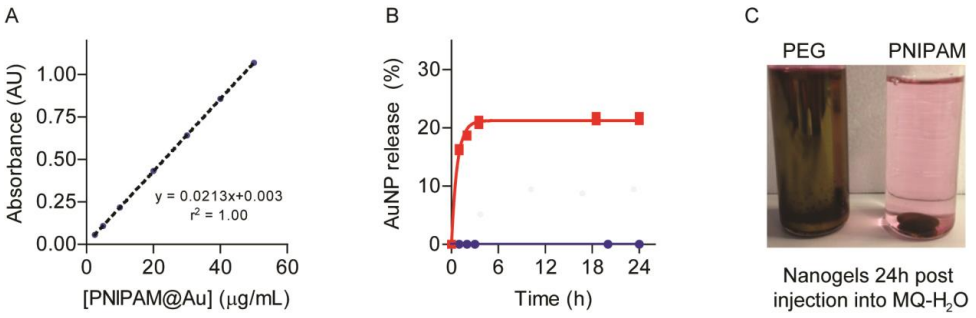


Figure S3. A) Standard curve of UV-absorbance at λ = 537 nm as a function of PNIPAM-coated AuNP concentration in MQ-H_2O (r^2 = 1.00); B) In vitro release kinetics from SAIB/EtOH/PLA (75:20:5) with either 30 mg mL^{-1} PEG-coated AuNPs or PNIPAM-coated AuNPs in MQ-H_2O at 37 °C. Each point represents the mean ± SEM (n = 3). C) Representative image of nanogels containing either PEG-coated or PNIPAM-coated AuNPs after 24 h incubation in MQ-H_2O.

3. X-ray Visualization of Nanogel in Mice/Canine

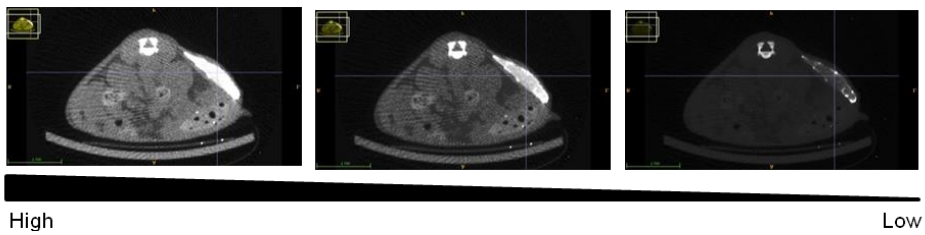


Figure S4. Visualization of nanogel composed of SAIB/EtOH/PLA (75:20:5) + 30 mg PNIPAM-coated AuNPs mL^{-1} administered subcutanously injection in immunocompetent mice. Aggregation of PNIPAM-coated AuNPs within the gel matrix visualized by reduction of the contrast level on the micro-CT image.

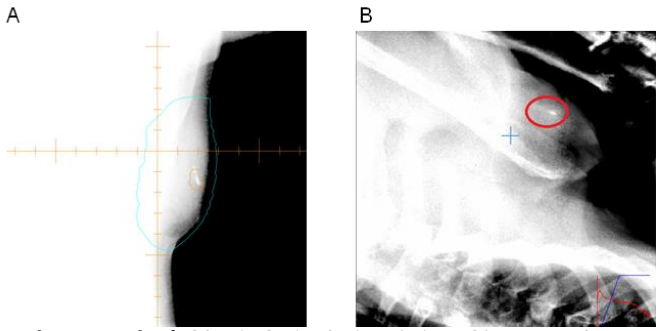


Figure S5. Nanogel composed of SAIB/EtOH/PLA (75:20:5) + 30mg PNIPAM-coated AuNPs mL^{-1} administered intratumorally into canine suffering from a solid subcutaneous mast cell tumor (maximum distance (x × y × z); 1.82 cm × 5.81 cm × 5.32 cm, CT volume; 31.64 cm^3) adherent to the underlying soft tissue located over the dorsal aspect of the left shoulder blade. A) OBI 2D x-ray imaging of nanogel and B) ExacTrac imaging of nanogel. Both images recorded prior to the first fraction of radiation therapy 6 days after injection.

Appendix B

Segmentation and estimation of the degradation rate over 14 weeks of two formulations of the x-SAIB gel in mice

NanoGuide project: Technical report

Segmentation and estimation of the degradation rate over 14 weeks of two formulations of the x-SAIB gel in mice

2nd edition

Anders Nymark Christensen, ANYM

December 1, 2015

1 Abstract

An experiment with 24 mice was started to test two formulations of the x-SAIB gel. The two formulations were injected into 8 mice each, with a third group of 8 mice as a control. The mice were scanned in a micro-CT scanner on day 0, 2, 8, 29, 51, and 98 after injection of the gel. The gel were semi-automatically segmented on the CT-scan using a Chan-Vese segmentation method. A Repeated Measurements Model (RMM) showed a non-significant difference in the volume loss of the gel over time between the two groups. The volume loss, with a 95% confidence interval, was $-0.09176\frac{\mu L}{d}, [-0.1134, -0.07009]\frac{\mu L}{d}$. This is less than what was found for the first 6 weeks.

Contents

1 Abstract **1**

2 Introduction **3**

3 Segmentation **3**

3.1 Chan-Vese method 3

3.2 Implementation 4

3.3 Volume results 6

3.4 Conclusion 6

4 Statistical analysis **7**

4.1 Model type 7

4.2 Model selection 7

4.3 Validation of model 10

4.4 Conclusion 10

5 Conclusion **13**

A Volume data **13**

B SAS program code **16**

C SAS output for chosen RMM model **21**

2 Introduction

An experiment with 24 mice in three groups (A, B and control) were initiated in the Cluster for Molecular Imaging at Copenhagen University on the 10th of July 2013. Groups A and B were given different formulations of a x-SAIB gel, while C acted as control. The mice were scanned in a micro-CT six times: on day 0, 2, 8, 29, 51 and 98 after injection of the gel. The goals of the experiment were multiple, but this report will focus narrowly on the stability of the gels, namely the rate with which they were degraded in the mice.

The segmentation process used to delineate the gel will be explained and possible errors from this process discussed. The statistical analysis of the volume data arising from the segmentation, will be explained in details and the degradation results for the different gels presented.

3 Segmentation

Given the CT scannings of the mice, we needed to isolate the gel in some way, in order to analyse the stability of the gels. As manual delineation is both time consuming and exhibits a large operator variability, an automatic approach was chosen.

The problem in image segmentation is to find some defining feature of the area one wants to segment. In our case the gels have a higher intensity than the surrounding tissue, and intensity is thus a good feature. Other features could be e.g. texture, shape or intensity distribution.

3.1 Chan-Vese method

The Chan-Vese method [1] works by supplying an initial guess on a segmentation, which will then be refined. The method has two components: An intensity component and a curvature component. The intensity component will try to align the segmentation along intensity edges in the image, i.e. gel tissue boundary in our case. The curvature component penalizes curvature, i.e. how curved or bended our segmentation is. This is a reasonable constraint as our gels mostly have a smooth surface. If we only looked at intensity, noise in the image could lead to a spiky segmentation and spurious angular edges.

The penalty function is given by

$$F(\phi) = \mu \int_\Omega |\nabla H(\phi(x,y))| \, dx \, dy + \lambda \int_\Omega |u(x,y) - c_1|^2 - |u(x,y) - c_2|^2 \, dx \, dy, \quad (1)$$

where H is the Heaviside function, $u(x,y)$ is the image, Ω describes the entire image, c_1 and c_2 are respectively the mean value outside and inside the level set, and μ and λ are weights.

The problem is how much weight to put on the first component μ vs. the second component λ. If the second component is to large we will cut of areas with high curvature, e.g. if the gel is flat and thin. If we weigh the second component

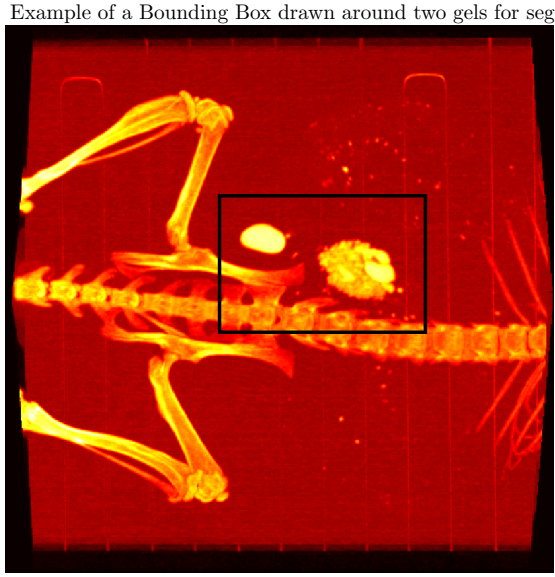
too little our segmentation will be too noisy. In this case the problem is not significant, as our gels have nice smooth borders, and the segmentation works for a wide range of parameters.

3.2 Implementation

A script for the segmentation was written in Matlab. The procedure is as follows. The folder containing the scans are selected, and one by one, the scans are loaded automatically into Matlab.

First a bounding box (BB) are drawn around the gel on Maximum Intensity Projections (MIP) in all three dimensions, as shown in fig. 1 for one of the dimensions.

Figure 1: Example of a Bounding Box drawn around two gels for segmentation

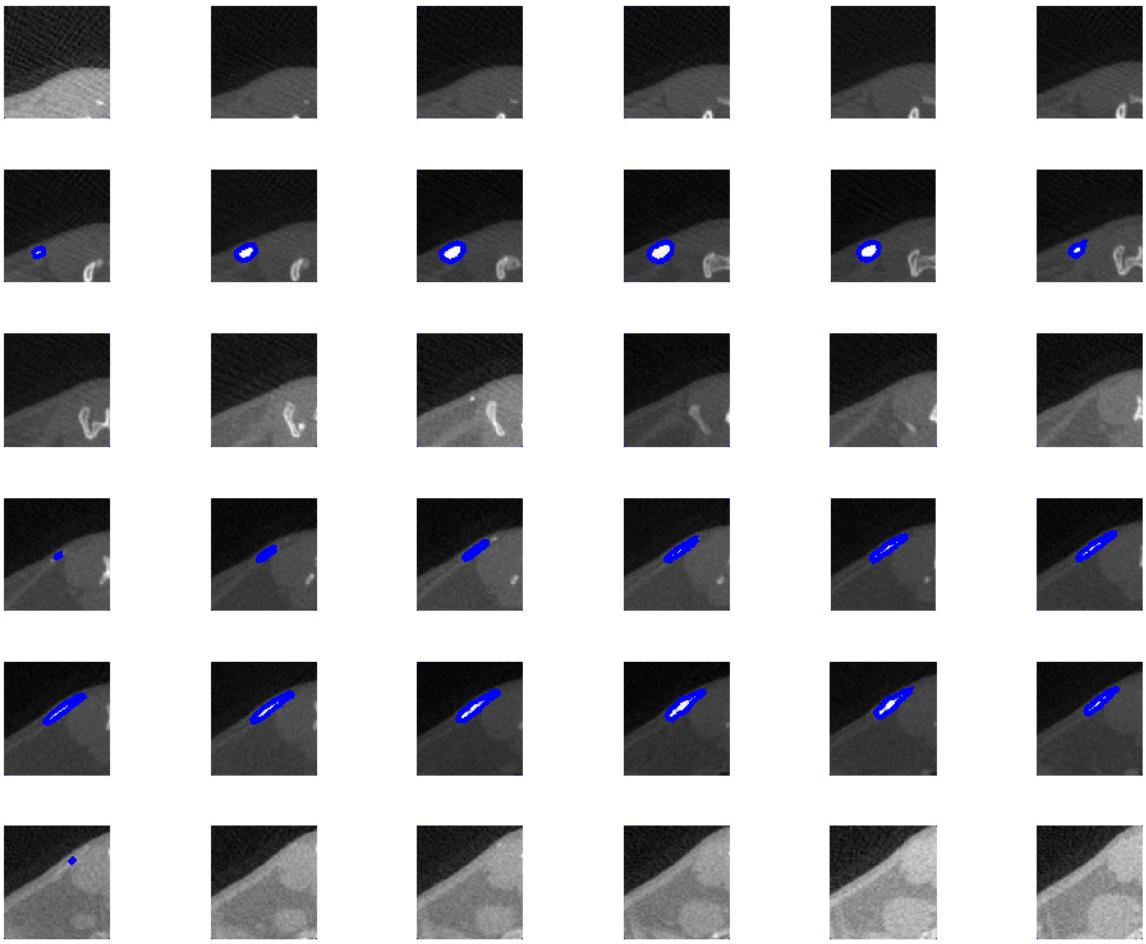


This BB was then automatically segmented into gel/non-gel by the Chan-Vese method, with the smoothing term $\mu = 0.25$ and the image term set to $\lambda = 5^{-6}$. The initial guess was obtained by finding the maximum value in the BB, subtracting 250, and then use that value as a threshold, assigning all voxels above that value as 'gel'. If the segmentation touched the edge of the BB that part of the segmentation was automatically removed. This was often the case for the spinal bones, as they have an intensity comparable to the gel.

The resulting segmentation was manually inspected on an image as seen in fig. 2. The bright white mass is the gel and the blue line the segmentation border.

If the segmentation is acceptable - i.e. if it follows the outline of the the gel on each slice - it is saved, otherwise it is possible to change the parameters in

Figure 2: Slices through two gels. The delineation is shown by blue. The images are individually scaled.

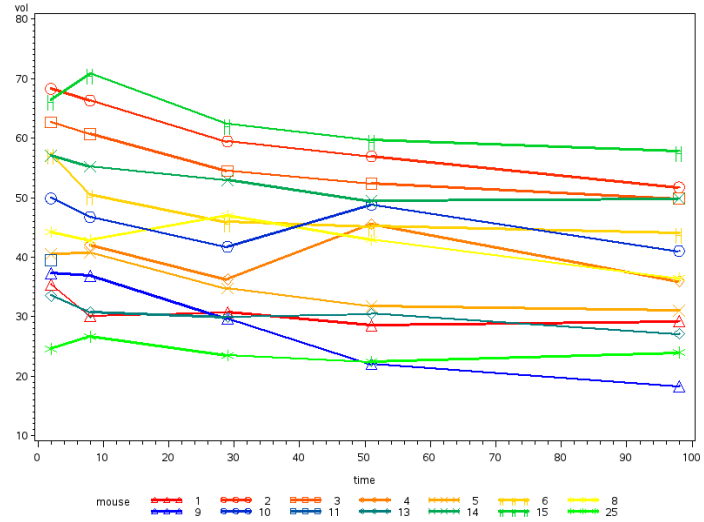


the model. This was necessary on a few scans where the gel was very flat and elongated. In those cases the curvature parameter was adjusted downwards, and a satisfactory segmentation achieved.

3.3 Volume results

The volumes for gel A and B can be seen in fig. 3 with the first measurement omitted. The reason for this, is that it mostly shows the diffusion of the ethanol from the gel and is not representative for the degradation of the gel. The volumes are listed in tabular form in Appendix A

Figure 3: Volume plotted over time. We see no apparent differences in the degradation rate of the two types. Group A: warm colours 1-8. Group B: cold colours, 9-25.



3.4 Conclusion

The gels can be segmented semi-automatically.

The segmentation is deterministic and we have no ground truth segmentation. It is therefore not sensible to assign a variance to the segmentation nor to the volume. If the segmentation for a gel is very dependant on the parameters chosen for the Chan-Vese model or very sensitive to a small perturbation of the image, it could be viewed as an "uncertain" segmentation. This degree of uncertainness might be incorporated into a statistical model. Not much research have been done is this area.

4 Statistical analysis

If we look at fig. 3 the curves looks mostly linear with a downwards trend. This is expected, as the gel is biodegradable. We will focus on this part, to see if the two gels have different properties regarding the degradation.

All data analysis was done in SAS or Matlab. The SAS program can be found in Appendix B.

4.1 Model type

As a first step, one might be tempted to fit a general linear model (GLM) (think linear regression with more than one variable). To capture the downwards trend, it is not far off, but for statistical tests it is inadequate. In GLM we assume that the measurements are uncorrelated. This is clearly not the case as we have repeated measurements on the same mice.

Instead we can use a repeated measurements model (RMM). Here we can model the correlation of the measurements made on the same mouse. This decreases the errors in the model, and increases the statistical test strength. The mice are still assumed uncorrelated.

The errors were modelled as autoregressive, giving the following variance-covariance matrix for a single mouse:

$$D(\epsilon_{jk}) = \sigma^2 \begin{bmatrix} 1 & \cdots & \rho^{|t(k)_{1j} - t(k)_{Ij}|} \\ \vdots & \ddots & \vdots \\ \rho^{|t(k)_{1j} - t(k)_{Ij}|} & \cdots & 1 \end{bmatrix} \tag{2}$$

where I is the last observation for mouse j, group k.

4.2 Model selection

We start by hypothesizing different models which may describe the data. We expect a different intercept for each mouse, as the amount of gel seems to vary, and we may also expect a mouse dependent slope, i.e. degradation rate, as each mouse may have different responses to the gel. A reasonable model may then be:

$$Y(k)_{ij} = \alpha(k)_j + \beta(k)_j \, t(k)_{ij} + \epsilon(k)_{ij} \tag{3}$$

where $Y(k)_{ij}$ are the volumes at time i for mouse j in group k, $\alpha(k)_j$ are the intercepts for mouse j in group k, $t(k)_{ij}$ are the times, $\beta(k)_j$ are slopes, and $\epsilon(k)_{ij}$ are the errors.

We will then successively test if a simpler model may provide a better explanation of the data, when taking into account that we use fewer parameters. A simpler model, that still adequately describes the data are preferred, as we have more degrees of freedom (DoF) left to account for the errors. The null hypothesis is thus, that the current model explains the data best against that a the previous model explains it just as well, when accounting for fewer free parameters. An overview of the tested models can found in table 1.

Table 1:

Model overview		
M1	$\alpha(k)_j + \beta(k)_j\, t(k)_{ij}$	Individual intercept and slope
M2	$\alpha(k)_j + \beta(k)\, t(k)_{ij}$	Individual intercept and group slope
M3	$\alpha(k)_j + \beta\, t(k)_{ij}$	Individual intercept and common slope
M4	$\alpha(k)_j$	Individual intercept
M5	$\alpha(k)$	Group intercept

This is done with the usual F-test statistic for the GLM. For the RMM we use that the -2 log-likelihood ratio is χ^2-distributed under the null hypothesis with DoF corresponding to the decrease in free parameters when moving to a simpler model. For the RMM we also use the corrected Akaike information criterion (AICC)[2], which weighs how good the model fits the data against how many parameters are used (smaller is better). The results are found for the GLM in table 2 and for the RMM in table 3. We have a total of 65 observations on 14 mice. Each group A and B have 7 mice.

Table 2:

GLM						
Model	Fixed effects	No. par.	RSS	DoF	F-statistic f	P(F >f)
M1	$\alpha(k)_j + \beta(k)_j\, t(k)_{ij}$	28	311.66	37		
M2	$\alpha(k)_j + \beta(k)\, t(k)_{ij}$	16	520.12	49	2.062347	0.0458
M3	$\alpha(k)_j + \beta\, t(k)_{ij}$	15	521.9696	50	0.174126	0.6783
M4	$\alpha(k)_j$	14	1208.80	51	32.43969	$1.06{\cdot}10^{-9}$
M5	$\alpha(k)$	2	10707.70	63	75.05399	$4.00{\cdot}10^{-28}$

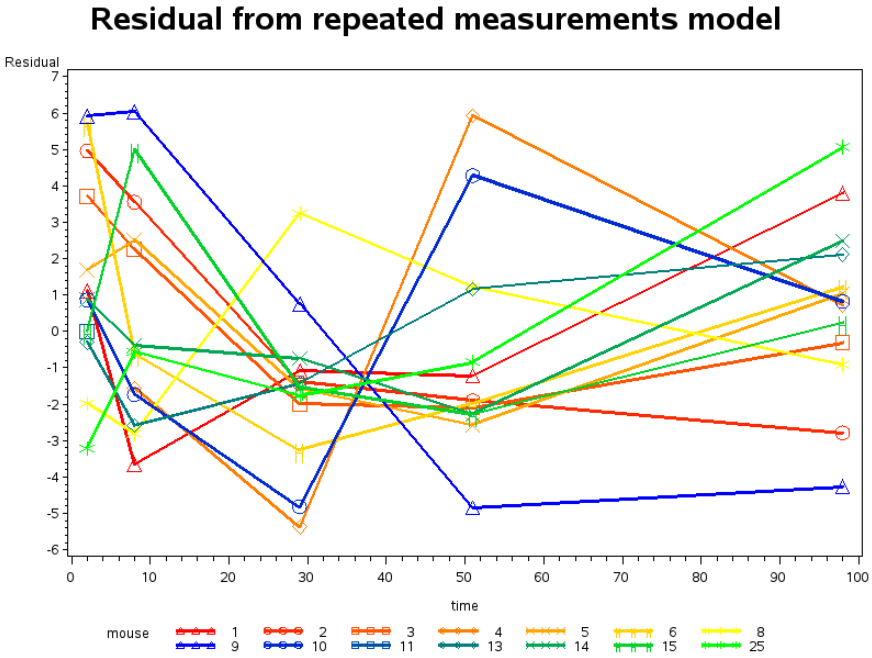
Table 3:

RMM							
Model	Fixed effects	No. par.	-2 log	Diff.	DoF	P(χ^2 >Diff.)	AICC
M1	$\alpha(k)_j + \beta(k)_j\, t(k)_{ij}$	28	286.40				387.5
M2	$\alpha(k)_j + \beta(k)\, t(k)_{ij}$	16	316.70	30.30	12	$2.516 \cdot 10^{-3}$	367.6
M3	$\alpha(k)_j + \beta\, t(k)_{ij}$	15	317	0.30	1	0.5839	364.0
M4	$\alpha(k)_j$	14	362.10	45.10	1	$1.87{\cdot}10^{-11}$	405.4
M6	$\alpha(k)$	2	409.40	4730	12	$4.14{\cdot}10^{-6}$	418.1

We see from table 2 and 3, that the step from M1 to M2 is significant at a 5% level. Taking statistical methodology very literally, we should thus use the full model M1. But obviously the superior fit to the model M1 is due to a very large number of parameters causing 'overfitting'. In model selection AICC is counter balancing the fit with the number of parameters used, and using AICC we should choose M3. To check whether M3 adequately describes the data we look at the residuals, which are shown in figure 4. As the residuals look random and as we can see no trend, we choose to accept M3 as an adequate model for the data. The full SAS-output for the model can be seen in appendix C.

We can thus conclude, that there is no significant difference between group A and B, as there is no significant difference between M2 where we have group

Figure 4: First line belong to group A (warm colours), second line to group B (cold colours)



slopes and M3 where we have a common slope. Using the RMM we find the slope, i.e. degradation rate, with a 95% confidence interval, to be:

$$\text{Degradation rate} = -0.09176\frac{\mu L}{d}, [-0.1134, -0.07009]\frac{\mu L}{d} \tag{4}$$

Based on the likelihood distance 2 observations were found to have a large influence. This was observation 18 and 65. Observation 18 is mouse 4 at day 51 and observation 65 is mouse 25 at day 98 which is the last for mouse 10. They are both clearly the result of an unexpected and physically unexplainable behaviour: That the volume of the gel increases. A full set of influence diagnostics can be found in appendix B.

4.3 Validation of model

An assumption we made, was that the volume could be described by a linear model. To test whether this is a reasonable assumption, we construct a physics model for the gel.

If we assume a that the gel is spherical and is broken down on the surface only, we can make the following model of the gel volume:

$$V(t) = \frac{4}{3}\pi(r - d \cdot t)^3 \tag{5}$$

$V(t)$ is the time dependant volume, r is the initial radius and d is the distance from the surface that is broken down per time step t.

Using the same equation we find that a gel with a volume of 50 μL has a radius of 2.2854 mm. With a degradation rate of $-0.09176\frac{\mu L}{d}$ as found, this corresponds to 0.0014 mm of the surface being broken down each day.

This is less than found for the first 51 days alone. The rate estimate there was $-0.1337\frac{\mu L}{d}$ corresponding to 0.0020 mm surface per day.

How this affects the volume is plotted in fig. 5. The time to complete degradation with this very simplistic model is more than a year longer. We see that the initial part is quite linear, at least for the first 100 days. The assumption of linear volume loss is thus reasonable, even though the gels are not exactly spherical.

If we use the upper and lower bounds on the 95 % confidence interval, we get the result shown in fig. 6.

4.4 Conclusion

A repeated measurements model was used and the specific model selected was performed based on AICC. The residuals were evaluated and showed no trend or organised structure.

Gel A and B experience the same degradation rate, which, with a 95% confidence interval, was found to be $-0.09176\frac{\mu L}{d}, [-0.1134, -0.07009]\frac{\mu L}{d}$.

Based on a physics model, the assumption of linear volume loss would seem a reasonable model.

Figure 5: Degradation of the gel assuming loss from the surface only, mean estimate.

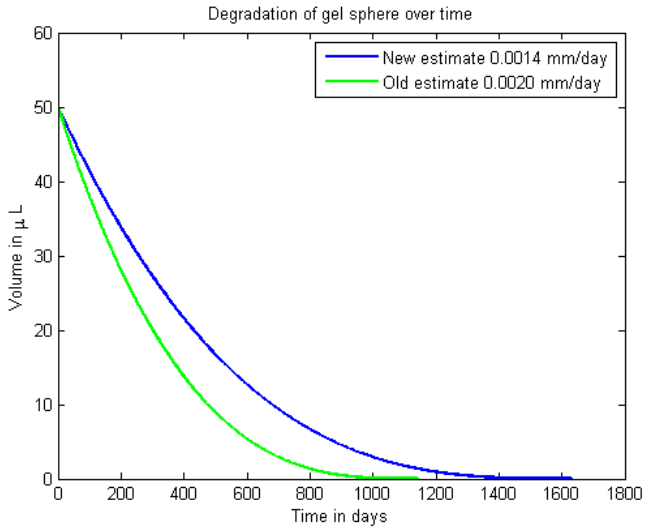
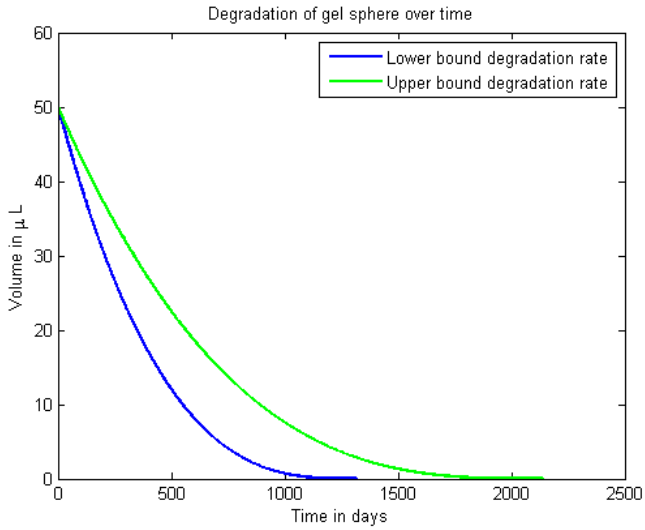


Figure 6: Degradation of the gel assuming loss from the surface only, upper and lower bound



5 Conclusion

The x-SAIB gels can be semi-automatically segmented, using a Chan-Vese segmentation model. More research is needed before segmentation 'uncertainty' can be quantified and incorporated into a statistical model with only a single measurement per time.

Gel A and B experience the same degradation rate, which, with a 95% confidence interval, was found to be $-0.09176\frac{\mu L}{d}$, $[-0.1134, -0.07009]\frac{\mu L}{d}$. A more robust estimate would require more mice.

Based on a physics model, the assumption of linear volume loss would seem a reasonable model.

References

[1] T. F. Chan and L. A. Vese, "Active contours without edges.," *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, vol. 10, pp. 266–277, Feb. 2001.

[2] T. Hurvich, "Bias of the corrected aic criterion for underfitted regression and time series models," *Biometrika*, vol. 76, pp. 297–307., 1989.

A Volume data

NanoGuide - gel volume data from mice

Obs	mouse	group	time	vol
1	1	1	2	35.4257
2	1	1	8	30.0960
3	1	1	29	30.7415
4	1	1	51	28.5923
5	2	1	2	68.3343
6	2	1	8	66.3791
7	2	1	29	59.5002
8	2	1	51	56.9742
9	3	1	2	62.7148
10	3	1	8	60.7109
11	3	1	29	54.5424
12	3	1	51	52.4118
13	4	1	8	42.0277
14	4	1	29	36.2904
15	4	1	51	45.5662
16	5	1	2	40.5240
17	5	1	8	40.7976
18	5	1	29	34.7843
19	5	1	51	31.7808
20	6	1	2	57.4094
21	6	1	8	50.5330
22	6	1	29	45.9641
23	6	1	51	45.2228
24	8	1	2	44.2484
25	8	1	8	42.8600
26	8	1	29	46.9823
27	8	1	51	42.9485
28	9	2	2	37.3119
29	9	2	8	36.8905
30	9	2	29	29.6868
31	9	2	51	22.0560
32	10	2	2	49.9240
33	10	2	8	46.7736
34	10	2	29	41.7671

NanoGuide - gel volume data from mice

Obs	mouse	group	time	vol
35	10	2	51	48.8465
36	11	2	2	39.5342
37	11	2	8	38.0792
38	13	2	2	33.5574
39	13	2	8	30.7204
40	13	2	29	29.9555
41	13	2	51	30.5272
42	14	2	2	57.0920
43	14	2	8	55.2756
44	14	2	29	53.0062
45	14	2	51	49.4588
46	15	2	2	66.4628
47	15	2	8	70.8847
48	15	2	29	62.4411
49	15	2	51	59.6805
50	25	2	2	24.5820
51	25	2	8	26.6777
52	25	2	29	23.5443
53	25	2	51	22.4587

B SAS program code

```
1
2  Data Nano2;
3  input mouse group time vol intensity variance surface;
4  cards;
5  1 1 2 35.42569864 -49.52979601 108669.995 80.04514464
6  1 1 8 30.09601067 51.44175255 106357.0777 78.08688025
7  1 1 29 30.74151676 43.37341856 125521.3162 80.07125483
8  1 1 51 28.59226568 77.53876299 129834.32 69.339966
9  1 1 98 29.29217291 37.0853476 134306.9352 78.07817686
10 2 1 2 68.33432953 109.4608959 109445.8243 122.3784689
11 2 1 8 66.37913626 142.7016342 119861.0259 126.4081419
12 2 1 29 59.5002337 281.1633461 105998.7397 115.9640651
13 2 1 51 56.97423442 275.7991136 117661.6498 91.82084114
14 2 1 98 51.78176725 357.1379716 116923.7708 86.79027752
15 3 1 2 62.7147728 274.0531597 122797.1628 107.6610242
16 3 1 8 60.71086209 303.5634404 110156.1762 96.65122662
17 3 1 29 54.54242219 404.2452735 99903.43663 75.37142031
18 3 1 51 52.41184614 474.9968397 100473.2707 67.62539674
19 3 1 98 49.89640231 504.5179164 86544.34748 61.23710315
20 4 1 8 42.02772375 266.8248102 126420.5463 84.66664858
21 4 1 29 36.2904332 334.9333035 98390.20222 57.98203257
22 4 1 51 45.56623383 225.8813949 187530.8619 72.1946803
23 4 1 98 36.04278621 389.7182248 120846.7979 56.80707394
24 5 1 2 40.52397876 236.1339438 111813.5532 87.66932064
25 5 1 8 40.79760838 219.1499224 120830.9048 82.02951921
26 5 1 29 34.78425234 366.0769374 107422.8303 54.37882611
27 5 1 51 31.78082215 448.4554304 90005.30525 45.25766577
28 5 1 98 31.06792361 446.5332044 91966.81657 42.59442621
29 6 1 2 57.40944355 217.1534121 117999.5588 101.3510611
30 6 1 8 50.53297686 343.3837008 106286.1012 89.94961072
31 6 1 29 45.96409292 427.1831688 90947.86316 60.39287361
32 6 1 51 45.22277587 453.8050129 97421.79974 54.81399597
33 6 1 98 44.11770194 488.9778228 89563.44952 56.39801427
34 8 1 2 44.24842707 238.3314739 111928.0403 95.09331851
35 8 1 8 42.85998002 289.1258857 114325.2922 85.38032716
36 8 1 29 46.98228742 185.7542644 148668.5902 84.13574135
37 8 1 51 42.94848337 246.7547405 145413.9902 72.57762978
38 8 1 98 36.50966168 418.1167797 114538.3369 65.51047121
39 9 2 2 37.31187553 1114.62812 346953.9974 89.34907631
40 9 2 8 36.89046967 1149.053264 386614.6842 88.39170261
41 9 2 29 29.68678418 547.2069088 588556.8859 77.13820995
42 9 2 51 22.05600912 779.5914077 433365.0605 49.33085569
43 9 2 98 18.33805647 996.5273412 459672.9202 43.34291837
44 10 2 2 49.92400886 853.9556647 324428.4407 93.70947834
45 10 2 8 46.77361439 1137.978891 356115.3594 93.33523226
46 10 2 29 41.76708544 1338.703927 302387.6231 65.30158967
47 10 2 51 48.84654147 932.6099835 344337.7039 62.89074863
```

```
10 2 98 41.07123801 1293.54629 333557.0259 57.45982874
11 2 2 39.53420277 1140.504292 307094.3966
11 2 8 38.07917522 1213.068724 332007.4716
13 2 2 33.5573848 1055.388758 327783.165 80.7849334
13 2 8 30.72040587 1254.557949 339645.6103 67.86909186
13 2 29 29.95554205 1229.502914 338112.0166 67.7472443
13 2 51 30.52716002 1204.186531 403008.8158 65.24066589
13 2 98 27.15266533 1478.579977 400421.6488 58.94810967
14 2 2 57.09196823 1161.750846 257795.559 87.26026097
14 2 8 55.27561967 1225.739413 282794.6129 88.52225357
14 2 29 53.00619891 1320.464845 269595.8769 82.83893516
14 2 51 49.4587573 1541.95556 286752.8728 73.27390157
14 2 98 49.89071861 1447.94906 297926.6473 73.70907143
15 2 2 66.46276787 1241.89777 281731.4135 124.0756314
15 2 8 70.88468753 1101.914434 350269.1201 125.6596497
15 2 29 62.44114317 1279.102169 321297.2344 101.107366
15 2 51 59.68048822 1385.244075 314925.7126 94.21427538
15 2 98 57.88768642 1409.830771 306761.2787 87.86079538
25 2 2 24.58200839 1241.273658 323114.2292 48.70421109
25 2 8 26.67767028 1091.256361 380429.9504 46.20633607
25 2 29 23.54432692 1325.256544 356743.8621 41.85463744
25 2 51 22.45873996 1459.392589 336672.244 37.01554856
25 2 98 24.05910787 1301.038811 427304.5996 37.51164221
;
ods graphics on;
ods pdf file="NanoGuide.pdf";

title 'NanoGuide - gel volume data from mice';
proc print;
run;

/* The following lines plot the data */;
title 'Gel volume from all mice';
filename outgraph 'VolData.png';
goptions reset=all device=png GSFNAME=outgraph;
symbol1 interpol=join width=2 value=triangle h=2 c=
    CXFF0000;
symbol2 interpol=join width=2 value=circle h=2 c=CXFF2A00
    ;
symbol3 interpol=join width=2 value=square h=2 c=CXFF5400
    ;
symbol4 interpol=join width=2 value=diamond h=2 c=
    CXFF7E00;
symbol5 interpol=join width=2 value=x h=2 c=CXFFA800;
symbol6 interpol=join width=2 value=hexagon h=2 c=
    CXFFD200;
symbol8 interpol=join width=2 value=star h=2 c=CXFFFF00;

symbol9 interpol=join width=2 value=triangle h=2 c=
    CX0000FF;
```

```
92  symbol10 interpol=join width=2 value=circle h=2 c=
        CX002AD2;
93  symbol11 interpol=join width=2 value=square h=2 c=
        CX0054A8;
94  symbol13 interpol=join width=2 value=diamond h=2 c=
        CX007E7E;
95  symbol14 interpol=join width=2 value=x h=2 c=CX00A854;
96  symbol15 interpol=join width=2 value=hexagon h=2 c=
        CX00D22A;
97  symbol25 interpol=join width=2 value=star h=2 c=CX00FF00;
98  Proc gplot data = Nano2;
99  Plot vol*time = mouse /name="VolData";
100 Run;
101
102 /*----------GLM----------*/;
103
104 /* Individual GLM model */;
105 title 'GLM - Individual slope and intercept';
106 proc glm data=Nano2 plots=all;
107 class mouse;
108 model vol = mouse time(mouse) /solution noint;
109 run;
110
111 /* Group slope individual intercept GLM model */;
112 title 'GLM - Group slope and individual intercept';
113 proc glm data=Nano2 plots=all;
114 class group mouse;
115 model vol = mouse time(group) /solution noint;
116 run;
117
118 /* One slope individual intercept GLM model */;
119 title 'GLM - Common slope and individual intercept';
120 proc glm data=Nano2 plots=all;
121 class group mouse;
122 model vol = mouse time /solution noint;
123 output out=resGLM r=resid;
124 run;
125
126 /* Individual intercept GLM model */;
127 title 'GLM - individual intercept';
128 proc glm data=Nano2 plots=all;
129 class mouse;
130 model vol = mouse /solution noint;
131 run;
132
133 /* Group slope and intercept*/;
134 title 'GLM - Group slope and intercept';
135 proc glm data=Nano2 plots=all;
136 class group;
137 model vol = group time(group) /solution noint;
```

```
138 run;
139
140 /* Group intercept */;
141 title 'Group intercept';
142 proc glm data=Nano2 plots=all;
143 class group;
144 model vol = group /solution noint;
145 run;
146
147 /*----------RMM----------*/;
148
149 /* Individual RMM model*/;
150 title 'RMM - Individual slope and intercept';
151 proc mixed data=Nano2 method=ml covtest plots=
      RESIDUALPANEL;
152 class mouse;
153 model vol = mouse time(mouse) /s outp=pred noint;
154 repeated / type=SP(POW)(time) subject=mouse rcorr;
155 run;
156
157 /*Group slope individual intercept*/;
158 title 'RMM - Group slope and individual intercept';
159 proc mixed data=Nano2 method=ml covtest plots=
      RESIDUALPANEL;
160 class group mouse;
161 model vol = mouse time(group) /s outp=pred noint;
162 repeated / type=SP(POW)(time) subject=mouse rcorr;
163 run;
164
165 /*One slope individual intercept*/;
166 title 'RMM - Common slope and individual intercept';
167 proc mixed data=Nano2 method=ml covtest plots=
      RESIDUALPANEL cl;
168 class group mouse;
169 model vol = mouse time /s outp=resRMM noint influence
      alpha=0.05 alphap=0.05;
170 repeated / type=SP(POW)(time) subject=mouse rcorr;
171 run;
172
173 /* Individual intercept*/;
174 title 'RMM - Individual intercept';
175 proc mixed data=Nano2 method=ml covtest plots=
      RESIDUALPANEL;
176 class mouse;
177 model vol = mouse /s outp=pred noint;
178 repeated / type=SP(POW)(time) subject=mouse rcorr;
179 run;
180
181 /*Group slope and intercept*/;
182 title 'RMM - Group slope and intercept';
```

```
183 proc mixed data=Nano2 method=ml covtest plots=
        RESIDUALPANEL;
184 class group;
185 model vol = group time(group) /s outp=pred noint;
186 repeated / type=SP(POW)(time) subject=mouse rcorr;
187 run;
188
189 /* Group intercept */;
190 title 'RMM - Group intercept';
191 proc mixed data=Nano2 method=ml covtest plots=
        RESIDUALPANEL;
192 class group;
193 model vol = group /s outp=pred noint;
194 repeated / type=SP(POW)(time) subject=mouse rcorr;
195 run;
196
197 filename outgraph 'ResData.png';
198 goptions reset=all device=png GSFNAME=outgraph;
199 symbol1 interpol=join width=2 value=triangle h=2 c=
        CXFF0000;
200 symbol2 interpol=join width=2 value=circle h=2 c=CXFF2A00
        ;
201 symbol3 interpol=join width=2 value=square h=2 c=CXFF5400
        ;
202 symbol4 interpol=join width=2 value=diamond h=2 c=
        CXFF7E00;
203 symbol5 interpol=join width=2 value=x h=2 c=CXFFA800;
204 symbol6 interpol=join width=2 value=hexagon h=2 c=
        CXFFD200;
205 symbol8 interpol=join width=2 value=star h=2 c=CXFFFF00;
206
207 symbol9 interpol=join width=2 value=triangle h=2 c=
        CX0000FF;
208 symbol10 interpol=join width=2 value=circle h=2 c=
        CX002AD2;
209 symbol11 interpol=join width=2 value=square h=2 c=
        CX0054A8;
210 symbol13 interpol=join width=2 value=diamond h=2 c=
        CX007E7E;
211 symbol14 interpol=join width=2 value=x h=2 c=CX00A854;
212 symbol15 interpol=join width=2 value=hexagon h=2 c=
        CX00D22A;
213 symbol25 interpol=join width=2 value=star h=2 c=CX00FF00;
214
215 Title 'Residual from repeated measurements model';
216 proc gplot data = resRMM;
217 plot resid*time = mouse ;
218 run;
219
220 /*
```

```
goptions reset=all;
symbol1 interpol=join width=4 value=triangle h=2 c=
    CXFF0000;
symbol2 interpol=join width=4 value=circle h=2 c=CXFF2A00
    ;
symbol3 interpol=join width=4 value=square h=2 c=CXFF5400
    ;
symbol4 interpol=join width=4 value=diamond h=2 c=
    CXFF7E00;
symbol5 interpol=join width=4 value=x h=2 c=CXFFA800;
symbol6 interpol=join width=4 value=hexagon h=2 c=
    CXFFD200;
symbol8 interpol=join width=4 value=star h=2 c=CXFFFF00;

symbol9 interpol=join width=4 value=triangle h=2 c=
    CX0000FF;
symbol10 interpol=join width=4 value=circle h=2 c=
    CX002AFF;
symbol11 interpol=join width=4 value=square h=2 c=
    CX0054FF;
symbol13 interpol=join width=4 value=diamond h=2 c=
    CX007EFF;
symbol14 interpol=join width=4 value=x h=2 c=CX00A8FF;
symbol15 interpol=join width=4 value=hexagon h=2 c=
    CX00D2FF;
symbol25 interpol=join width=4 value=star h=2 c=CX00FFFF;

Title 'Plotting with fitted regression lines';
proc gplot data = resRMM;
plot resid*time = mouse;
run;
*/;

ods pdf close;
ods graphics off;
```

C SAS output for chosen RMM model

RMM - Common slope and individual intercept

The Mixed Procedure

Model Information	
Data Set	WORK.NANO2
Dependent Variable	vol
Covariance Structure	Spatial Power
Subject Effect	mouse
Estimation Method	ML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Between-Within

Class Level Information		
Class	**Levels**	**Values**
group	2	1 2
mouse	14	1 2 3 4 5 6 8 9 10 11 13 14 15 25

Dimensions	
Covariance Parameters	2
Columns in X	15
Columns in Z	0
Subjects	14
Max Obs Per Subject	5

Number of Observations	
Number of Observations Read	65
Number of Observations Used	65
Number of Observations Not Used	0

Iteration History			
Iteration	**Evaluations**	**-2 Log Like**	**Criterion**
0	1	319.87143872	
1	2	317.06209400	0.00068515
2	1	316.98605809	0.00003302

RMM - Common slope and individual intercept

The Mixed Procedure

Iteration History			
Iteration	**Evaluations**	**-2 Log Like**	**Criterion**
3	1	316.98267987	0.00000010
4	1	316.98266951	0.00000000

Convergence criteria met.

Estimated R Correlation Matrix for mouse 1					
Row	**Col1**	**Col2**	**Col3**	**Col4**	**Col5**
1	1.0000	0.4116	0.01840	0.000710	6.774E-7
2	0.4116	1.0000	0.04472	0.001725	1.646E-6
3	0.01840	0.04472	1.0000	0.03857	0.000037
4	0.000710	0.001725	0.03857	1.0000	0.000954
5	6.774E-7	1.646E-6	0.000037	0.000954	1.0000

Covariance Parameter Estimates								
Cov Parm	**Subject**	**Estimate**	**Standard Error**	**Z Value**	**Pr Z**	**Alpha**	**Lower**	**Upper**
SP(POW)	mouse	0.8625	0.06262	13.77	<.0001	0.05	0.7397	0.9852
Residual		7.9545	1.4228	5.59	<.0001	0.05	5.7649	11.6868

Fit Statistics	
-2 Log Likelihood	317.0
AIC (smaller is better)	351.0
AICC (smaller is better)	364.0
BIC (smaller is better)	361.8

Null Model Likelihood Ratio Test		
DF	**Chi-Square**	**Pr > ChiSq**
1	2.89	0.0892

RMM - Common slope and individual intercept

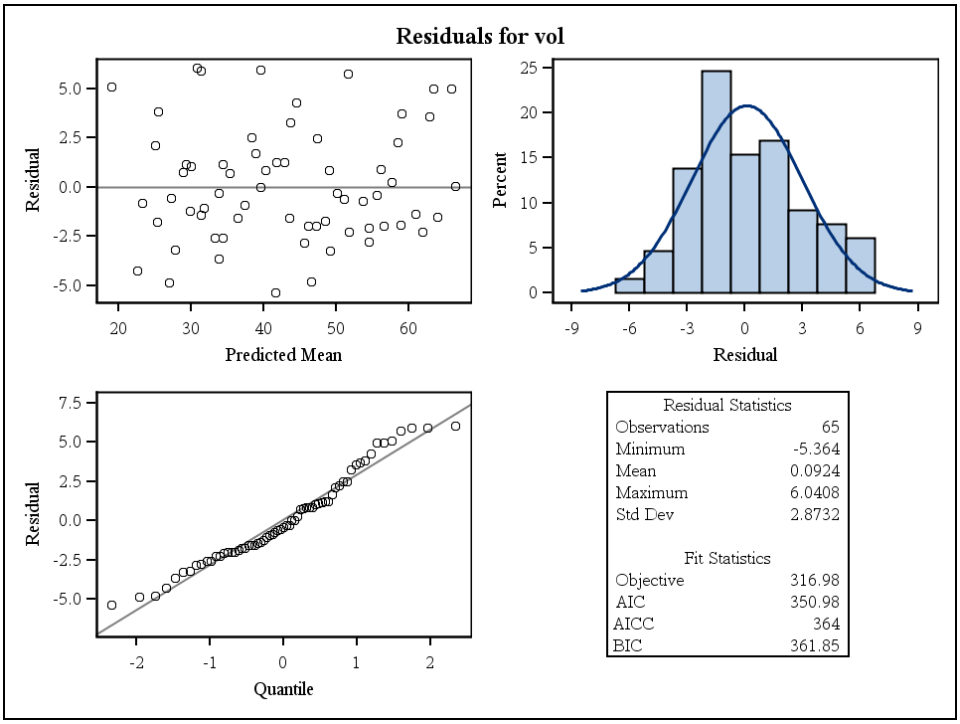
The Mixed Procedure

Solution for Fixed Effects									
Effect	**mouse**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**	**Alpha**	**Lower**	**Upper**
mouse	1	34.4845	1.4417	0	23.92	.	0.05	.	.
mouse	2	63.5541	1.4417	0	44.08	.	0.05	.	.
mouse	3	59.1857	1.4417	0	41.05	.	0.05	.	.
mouse	4	44.3160	1.5271	0	29.02	.	0.05	.	.
mouse	5	39.0192	1.4417	0	27.06	.	0.05	.	.
mouse	6	51.8735	1.4417	0	35.98	.	0.05	.	.
mouse	8	46.4017	1.4417	0	32.18	.	0.05	.	.
mouse	9	31.5838	1.4417	0	21.91	.	0.05	.	.
mouse	10	49.2355	1.4417	0	34.15	.	0.05	.	.
mouse	11	39.7177	2.8205	0	14.08	.	0.05	.	.
mouse	13	34.0362	1.4417	0	23.61	.	0.05	.	.
mouse	14	56.3982	1.4417	0	39.12	.	0.05	.	.
mouse	15	66.6286	1.4417	0	46.21	.	0.05	.	.
mouse	25	27.9729	1.4417	0	19.40	.	0.05	.	.
time		-0.09176	0.01079	50	-8.51	<.0001	0.05	-0.1134	-0.07009

Type 3 Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
mouse	14	0	519.95	.
time	1	50	72.34	<.0001

RMM - Common slope and individual intercept

The Mixed Procedure



Residual Statistics	
Observations	65
Minimum	-5.364
Mean	0.0924
Maximum	6.0408
Std Dev	2.8732

Fit Statistics	
Objective	316.98
AIC	350.98
AICC	364
BIC	361.85

RMM - Common slope and individual intercept

The Mixed Procedure

Influence Diagnostics										
Deleted Obs. Index	Observed Value	Predicted Value	Residual	Leverage	PRESS Residual	Internally Studentized Residual	RMSE without deleted obs	Externally Studentized Residual	Cook's D	DFFITS
1	35.4257	34.301	1.1248	0.185	1.6717	0.4633	2.81651	0.4639	0.00969	0.38168
2	30.096	33.750	-3.6543	0.167	-4.4034	-1.4985	2.78103	-1.5197	0.01865	-0.53624
3	30.7415	31.823	-1.0818	0.219	-1.3202	-0.4393	2.83970	-0.4363	0.00200	-0.17223
4	28.5923	29.804	-1.2122	0.227	-1.5535	-0.4918	2.83746	-0.4888	0.00413	-0.24753
5	29.2922	25.492	3.8006	0.280	5.2788	1.5884	2.78661	1.6076	0.06534	1.00203
6	68.3343	63.371	4.9638	0.185	5.6918	2.0444	2.79643	2.0619	0.01717	0.51168
7	66.3791	62.820	3.5592	0.167	3.8480	1.4594	2.83330	1.4528	0.00277	0.20291
8	59.5002	60.893	-1.3927	0.219	-1.7969	-0.5655	2.83477	-0.5627	0.00576	-0.29254
9	56.9742	58.874	-1.8999	0.227	-2.4357	-0.7707	2.83033	-0.7680	0.01019	-0.38964
10	51.7818	54.561	-2.7794	0.280	-3.8594	-1.1616	2.81272	-1.1648	0.03488	-0.72530
11	62.7148	59.002	3.7126	0.185	4.2916	1.5291	2.81338	1.5329	0.01086	0.40449
12	60.7109	58.452	2.2593	0.167	2.4077	0.9264	2.83994	0.9200	0.00073	0.10403
13	54.5424	56.525	-1.9821	0.219	-2.5296	-0.8049	2.82846	-0.8026	0.01057	-0.39709
14	52.4118	54.506	-2.0939	0.227	-2.6802	-0.8494	2.82796	-0.8472	0.01220	-0.42670
15	49.8964	50.193	-0.2964	0.280	-0.4109	-0.1239	2.84199	-0.1229	0.00039	-0.07609
16	42.0277	43.582	-1.5542	0.271	-2.0363	-0.6508	2.83589	-0.6473	0.00688	-0.31953
17	36.2904	41.655	-5.3644	0.244	-7.1019	-2.2192	2.73176	-2.2912	0.09533	-1.23462
18	45.5662	39.636	5.9302	0.251	7.9583	2.4455	2.70231	2.5523	0.13222	1.46979
19	36.0428	35.323	0.7196	0.298	1.0226	0.3046	2.84033	0.3025	0.00258	0.19532
20	40.524	38.836	1.6883	0.185	1.8244	0.6953	2.84073	0.6904	0.00060	0.09418
21	40.7976	38.285	2.5125	0.167	2.8575	1.0302	2.82943	1.0269	0.00396	0.24274
22	34.7843	36.358	-1.5738	0.219	-2.0079	-0.6391	2.83361	-0.6361	0.00665	-0.31424
23	31.7808	34.339	-2.5584	0.227	-3.2847	-1.0379	2.82025	-1.0379	0.01873	-0.53001
24	31.0679	30.026	1.0416	0.280	1.4475	0.4353	2.83816	0.4326	0.00493	0.27017
25	57.4094	51.690	5.7195	0.185	6.9615	2.3556	2.70659	2.4547	0.04998	0.90187
26	50.533	51.139	-0.6064	0.167	-1.1275	-0.2486	2.81283	-0.2493	0.00902	-0.36882
27	45.9641	49.212	-3.2482	0.219	-4.1082	-1.3190	2.80799	-1.3249	0.02608	-0.62827
28	45.2228	47.194	-1.9707	0.227	-2.5075	-0.7995	2.83029	-0.7967	0.01023	-0.39030
29	44.1177	42.881	1.2371	0.280	1.7189	0.5170	2.83646	0.5141	0.00694	0.32082
30	44.2484	46.218	-1.9698	0.185	-2.1392	-0.8113	2.83986	-0.8057	0.00093	-0.11726

RMM - Common slope and individual intercept

The Mixed Procedure

Influence Diagnostics		
Deleted Obs. Index	COVRATIO	Likelihood Distance
1	1.0784	0.1461
2	0.7223	0.3136
3	1.5350	0.0357
4	1.5335	0.0660
5	0.9675	1.0231
6	0.8701	0.2716
7	1.2627	0.0439
8	1.4571	0.0889
9	1.4220	0.1534
10	1.2798	0.5270
11	1.0431	0.1646
12	1.3546	0.0170
13	1.3628	0.1587
14	1.3866	0.1830
15	1.7459	0.0133
16	1.5928	0.1060
17	0.4945	1.6596
18	0.3653	2.4050
19	1.7594	0.0446
20	1.3943	0.0156
21	1.2120	0.0603
22	1.4394	0.1016
23	1.2776	0.2809
24	1.6767	0.0781
25	0.3267	1.0407
26	1.0159	0.1370
27	1.0960	0.3972
28	1.4213	0.1539
29	1.6468	0.1071
30	1.3815	0.0199

RMM - Common slope and individual intercept

The Mixed Procedure

Influence Diagnostics										
Deleted Obs. Index	Observed Value	Predicted Value	Residual	Leverage	PRESS Residual	Internally Studentized Residual	RMSE without deleted obs	Externally Studentized Residual	Cook's D	DFFITS
31	42.86	45.668	-2.8076	0.167	-3.1970	-1.1513	2.82589	-1.1490	0.00504	-0.27430
32	46.9823	43.741	3.2417	0.219	4.1495	1.3164	2.80404	1.3241	0.02906	0.66408
33	42.9485	41.722	1.2267	0.227	1.5472	0.4977	2.83804	0.4946	0.00365	0.23240
34	36.5097	37.409	-0.8992	0.280	-1.2492	-0.3758	2.83923	-0.3733	0.00366	-0.23290
35	37.3119	31.400	5.9116	0.185	6.6244	2.4348	2.79835	2.4539	0.01646	0.50059
36	36.8905	30.850	6.0408	0.167	6.6993	2.4770	2.79507	2.4994	0.01441	0.46904
37	29.6868	28.923	0.7641	0.219	0.9503	0.3103	2.84072	0.3081	0.00122	0.13446
38	22.056	26.904	-4.8478	0.227	-6.2643	-1.9666	2.75742	-2.0115	0.07123	-1.05722
39	18.3381	22.591	-4.2529	0.280	-5.9047	-1.7774	2.77259	-1.8080	0.08159	-1.12537
40	49.924	49.052	0.8720	0.185	1.2014	0.3591	2.83299	0.3575	0.00352	0.22852
41	46.7736	48.501	-1.7278	0.167	-2.0783	-0.7085	2.82901	-0.7063	0.00408	-0.24667
42	41.7671	46.574	-4.8073	0.219	-6.1458	-1.9522	2.75842	-1.9960	0.06319	-0.99541
43	48.8465	44.556	4.2910	0.227	5.5919	1.7407	2.77088	1.7718	0.06008	0.96624
44	41.0712	40.243	0.8286	0.280	1.1492	0.3463	2.83973	0.3439	0.00307	0.21325
45	39.5342	39.534	0.0000	1.000	.	.	.	.	.	.
46	33.5574	33.853	-0.2953	0.185	-0.1356	-0.1216	2.84013	-0.1208	0.00083	0.11047
47	30.7204	33.302	-2.5817	0.167	-3.0241	-1.0586	2.82109	-1.0583	0.00651	-0.31222
48	29.9555	31.375	-1.4195	0.219	-1.7884	-0.5764	2.83603	-0.5732	0.00480	-0.26685
49	30.5272	29.356	1.1709	0.227	1.5266	0.4750	2.83704	0.4722	0.00449	0.25802
50	27.1527	25.043	2.1093	0.280	2.9291	0.8816	2.82531	0.8800	0.02009	0.54806
51	57.092	56.215	0.8773	0.185	1.0931	0.3613	2.83832	0.3590	0.00151	0.14943
52	55.2756	55.664	-0.3885	0.167	-0.5211	-0.1593	2.84042	-0.1582	0.00058	-0.09297
53	53.0062	53.737	-0.7309	0.219	-0.9022	-0.2968	2.84097	-0.2946	0.00104	-0.12370
54	49.4588	51.718	-2.2595	0.227	-2.9088	-0.9166	2.82470	-0.9152	0.01496	-0.47304
55	49.8907	47.405	2.4854	0.280	3.4526	1.0387	2.81861	1.0394	0.02798	0.64819
56	66.4628	66.445	0.0177	0.185	-0.4059	0.0073	2.82686	0.0073	0.00582	-0.29454
57	70.8847	65.894	4.9902	0.167	5.9164	2.0462	2.74808	2.1001	0.02850	0.67092
58	62.4411	63.967	-1.5263	0.219	-1.9806	-0.6198	2.83278	-0.6171	0.00728	-0.32904
59	59.6805	61.949	-2.2681	0.227	-2.9103	-0.9201	2.82508	-0.9186	0.01464	-0.46786
60	57.8877	57.636	0.2520	0.280	0.3508	0.1053	2.84208	0.1045	0.00029	0.06568

RMM - Common slope and individual intercept

The Mixed Procedure

Influence Diagnostics		
Deleted Obs. Index	COVRATIO	Likelihood Distance
31	1.1674	0.0758
32	1.0507	0.4454
33	1.5429	0.0591
34	1.6957	0.0600
35	0.8882	0.2588
36	0.8401	0.2307
37	1.5517	0.0248
38	0.6499	1.1850
39	0.8316	1.3049
40	1.2847	0.0548
41	1.2067	0.0621
42	0.6423	1.0559
43	0.7522	0.9749
44	1.7047	0.0515
45	.	.
46	1.3855	0.0185
47	1.1093	0.0975
48	1.4767	0.0752
49	1.5267	0.0711
50	1.4633	0.3008
51	1.3592	0.0276
52	1.3615	0.0151
53	1.5557	0.0221
54	1.3395	0.2241
55	1.3627	0.4202
56	1.2040	0.0875
57	0.5052	0.5395
58	1.4267	0.1108
59	1.3449	0.2193
60	1.7475	0.0119

RMM - Common slope and individual intercept

The Mixed Procedure

Influence Diagnostics										
Deleted Obs. Index	**Observed Value**	**Predicted Value**	**Residual**	**Leverage**	**PRESS Residual**	**Internally Studentized Residual**	**RMSE without deleted obs**	**Externally Studentized Residual**	**Cook's D**	**DFFITS**
61	24.582	27.789	-3.2074	0.185	-3.8267	-1.3210	2.80919	-1.3263	0.01243	-0.43329
62	26.6777	27.239	-0.5611	0.167	-0.4100	-0.2301	2.83985	-0.2285	0.00076	0.10594
63	23.5443	25.312	-1.7674	0.219	-2.2350	-0.7177	2.83221	-0.7147	0.00771	-0.33870
64	22.4587	23.293	-0.8342	0.227	-1.0583	-0.3384	2.84023	-0.3360	0.00178	-0.16236
65	24.0591	18.980	5.0791	0.280	7.0543	2.1227	2.74207	2.1833	0.11667	1.36066

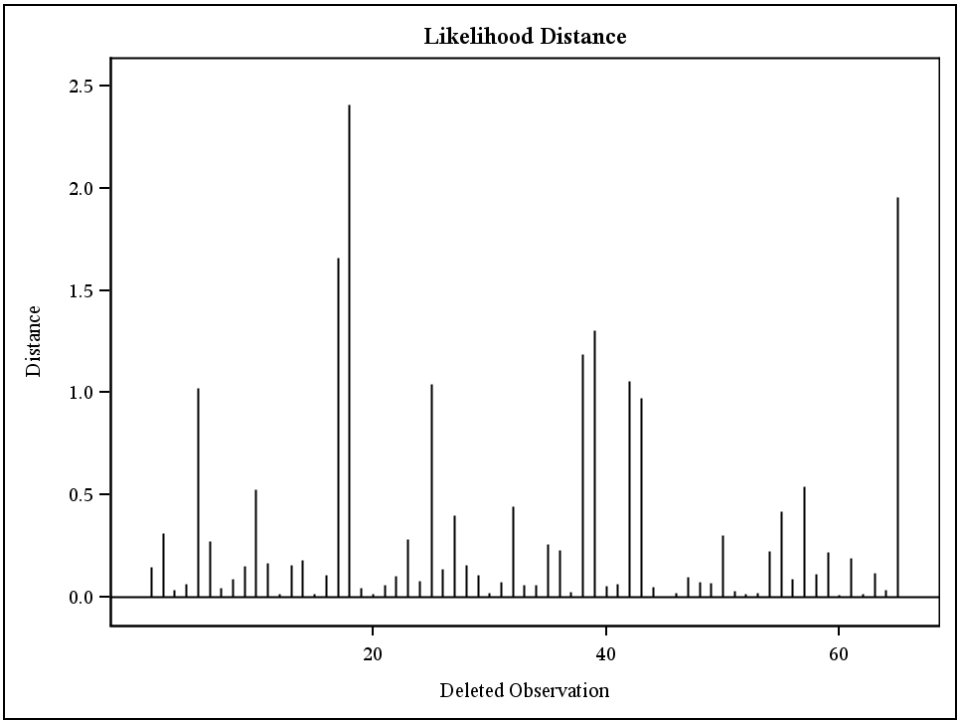
RMM - Common slope and individual intercept

The Mixed Procedure

Influence Diagnostics		
Deleted Obs. Index	**COVRATIO**	**Likelihood Distance**
61	0.9974	0.1900
62	1.3534	0.0174
63	1.4182	0.1170
64	1.5790	0.0327
65	0.5966	1.9564

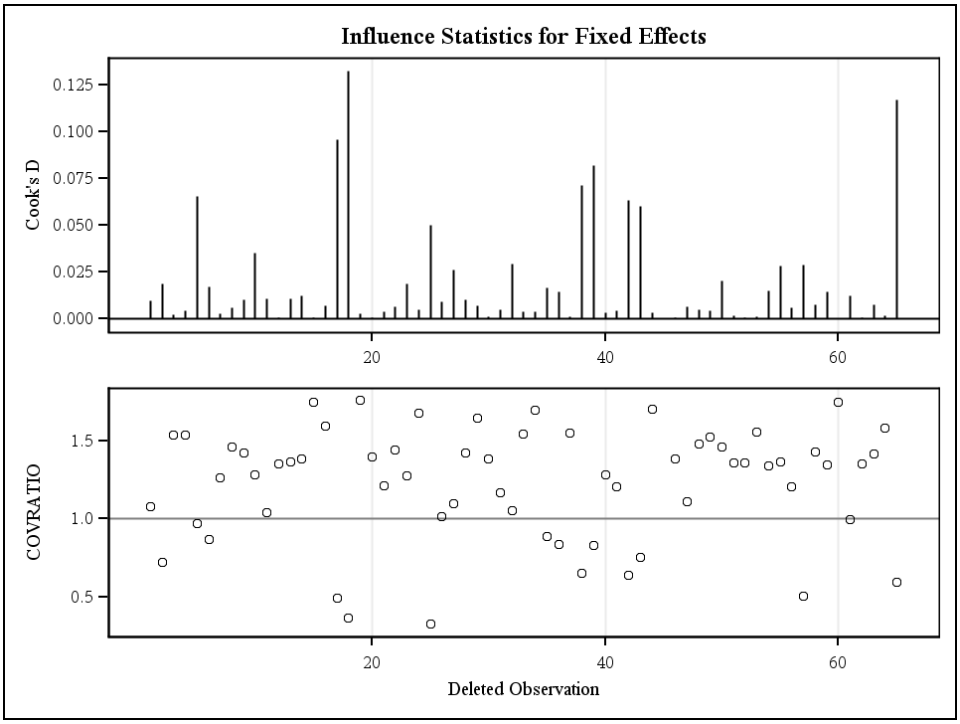
RMM - Common slope and individual intercept

The Mixed Procedure



RMM - Common slope and individual intercept

The Mixed Procedure



Appendix C

Injectable Silver Nanogels: In Vivo Dosimetry for External Beam Radiotherapy using Positron Emission Tomography

Injectable Silver Nanogels: *In Vivo* Dosimetry for External Beam Radiotherapy using Positron Emission Tomography

Anders N. Christensen[a], Jonas S. Rydhög[b], Rikke V. Søndergaard[c], Thomas L. Andresen[c], Søren Holm[d], Per Munck af Rosenschöld[b], Knut Conradsen[b], Rasmus I. Jølck[c, *]

[a] DTU Compute, Department of Applied Mathematics and Computer Science, Technical University of Denmark, Building 321/324, Matematiktorvet, 2800 Kgs. Lyngby, Denmark

[b] Department of Oncology, Section of Radiotherapy, 3994, Rigshospitalet, Blegdamsvej 9, 2100 Copenhagen, Denmark and The Niels Bohr Institute, University of Copenhagen, Blegdamsvej 17, 2100 Copenhagen, Denmark

[c] DTU Nanotech, Department of Micro- and Nanotechnology, Center for Nanomedicine and Theranostics, Technical University of Denmark, Building 423, 2800 Kgs. Lyngby, Denmark

[d] Department of Clinical Physiology, Nuclear Medicine and PET, Rigshospitalet, Blegdamsvej 9, 2100 Copenhagen, Denmark

* Corresponding Author; E-mail: rasj@nanotech.dtu.dk

Abstract: Development of safe and efficient radiotherapy routines requires quantification of the delivered absorbed dose to the cancer tissue in individual patients. *In vivo* dosimetry can provide accurate information about the absorbed dose delivered during treatment. In the current study, a novel silver-nanogel formulation based on poly(vinylpyrrolidinone)-coated silver nanoparticles formulated in a gelation matrix composed of sucrose acetate isobutyrate has been developed for use as an *in vivo* dosimeter for external beam radiotherapy. *In situ* photonuclear reactions trigger the formation of radioactive ^{106}Ag, which enables post treatment verification of delivered dose using positron emission tomography imaging. The silver-nanogel was investigated in a tissue equivalent thorax phantom using clinical settings and workflow for both standard fractionated radiotherapy (2 Gy) and stereotactic radiotherapy (10- and 22 Gy). The developed silver-nanogel provided high radiopacity on the planning CT-scans sufficient for marker position use in modern radiotherapy and provided dosimetric information about the absorbed dose with a 10% and 8% standard deviation for the stereotactic regimens, 10 and 22 Gy, respectively.

Keywords: Silver nanoparticles, *in vivo* dosimetry, external beam radiotherapy, positron emission tomography, sucrose acetate isobutyrate

The global burden of cancer continues to increase due to the increasing aging of the population and adoption of cancer-causing behaviors.[1] Cancer treatment includes surgery, radiation, chemotherapy, hormone therapy and immune therapy. Among these, radiotherapy is a cost-effective treatment modality and approximately 50% of all cancer patients receive radiotherapy either alone or in combination with other interventions.[2,3,4,5] Safe and efficient radiotherapy requires verification of the delivered absorbed dose to each individual patient[6,7,8] and is compulsory by law in many European countries[9]. Ideally, radiotherapy should be verified using a detector or a set of detectors that can be placed inside the patient, i.e. *in vivo*, and thus, provide relevant information of the absorbed dose delivered. Dosimetric readout should be feasible in a non-invasive manner and the dosimeter should be reusable, as radiotherapy is delivered in a series of smaller fractions over days to several weeks. Furthermore, placement of the detector itself should be as non-invasive as possible to ensure patient compliance.

Traditionally, *in vivo* dosimetry was commonly employed by using either silicon based diodes or radiothermoluminescent chips placed externally on the patient.[8] Silicon based diodes suffer from batch-to-batch variations,[8] drift in sensitivity over clinical lifetime,[10] temperature sensitivity[11] and require an electrical connection during irradiation. Radiothermoluminescent chips are based upon the ability of imperfect crystals to absorb and store the energy of ionizing radiation, which upon heating is re-emitted in the form of electromagnetic radiation, mainly in the visible part of the spectrum. The light emitted is then detected by a photomultiplier tube and the signal correlated to the absorbed dose. Radiothermoluminescent dosimeters placed internally have to be removed from the patient before readout is possible which limits its use in anatomical positions not readily accessible.[12] Following the clinical introduction of modern radiation delivery technology, such as intensity modulated volumetric arc therapy (VMAT),[13,14] the use of *in vivo* dosimeters has fallen into disuse in radiotherapy clinics. For radiation treatment using VMAT, dose measurements performed at the patient's skin surface cannot practically be used to estimate the radiation dose delivered to a deep-seated tumor or organ at risk. Further, intra cavitary placement of *in vivo* dosimeters is impractical for routine clinical use. This creates an urgent need for development of internal dosimeters, which can provide dosimetric information about the absorbed dose in a non-invasive manner.

Radiotherapy exploits the use of high-energy photon beams to induce DNA damage to cancerous tissue hereby killing or controlling tumor growth. High-energy photon beams can also be exploited to remove neutrons from the nucleus of stable elements by photonuclear reactions.[15,16] Neutron removal in certain elements results in the formation of radioactive isotopes, which can be detected using positron emission tomography (PET) imaging. In order to efficiently detect the radioactive decay using PET imaging, a positron-emitting radionuclide with sufficient half-life ($T_{½}$) must be formed during the photonuclear reaction. Photonuclear reactions using high-energy photon beams on various elements, including Ag, Cr, Cu, Ni, Sn, Ti and Zn, have been evaluated by Hansen *et. al*.[17] who concluded that only Ag, Cu and Zn, in which the positron-emitting radionuclides; ^{108}Ag, ^{106}Ag, ^{64}Cu ^{62}Cu ^{65}Zn and ^{63}Zn are being formed, resulted in detectable PET signals. Among these, Ag displayed the highest activity (28 Bq/mm^3) compared to Cu (17 Bq/mm^3) and Zn (1.1 Bq/mm^3), hence, less material was required for obtaining a detectable PET-signal.

We recently described the concept of an injectable liquid fiducial marker for use in image-guided radiotherapy (IGRT) by exploiting the use of polymer-coated gold nanoparticles.[18,19] In this present paper, we describe the proof-of-principle of a liquid injectable dosimetric silver nanogel marker,

which enables dual use as an IGRT fiducial marker as well as an internal *in vivo* dosimeter through readout of the radiation induced activation.[20,21] The silver nanogel is based on poly(vinylpyrrolidinone) (PVP) coated silver nanoparticles (AgNPs) in a nonsolvent-induced phase separation (NIPS) gelation matrix based on sucrose acetate isobutyrate (SAIB), further referred to as Ag-nanogel. SAIB, a fully acylated derivative of sucrose, is a biodegradable nonpolymeric compound, which has been approved for oral consumption and is generally regarded as safe (GRAS) by the FDA. SAIB mixtures with ethanol (EtOH) behaves as Newtonian liquids and have a low viscosity (50–200 mPa·s) suitable for injection with thin hypodermic needles.[22,23,24] Upon hydration, a highly viscous (gel-like) liquid is formed, which efficiently encapsulate high concentrations of PVP-coated AgNPs.

Results and Discussion

External radiotherapy relies on the ability to deliver high radiation doses to tumor cells with minimal exposure to surrounding tissues. Verification of the delivered absorbed dose to each individual patient is a requisite for developing new and safe treatment regimens, especially with the increasing focus on stereotactic radiation procedures. Here, we report an injectable Ag-nanogel for use as an internal *in vivo* dosimeter that can be placed intratumorally and detect therapeutic levels of ionizing radiation using PET for direct dosimetric readout. The Ag-nanogel have been evaluated in a tissue equivalent thorax phantom using clinical settings for both standard fractionated radiotherapy (dose of 2 Gy) and stereotactic radiotherapy (doses of 10 - and 22 Gy), which is often the standard treatment for inoperable cancers, with 33, 5 or 3 treatment fractions, respectively. Based on *in situ* induced photonuclear reactions the Ag-nanogel enabled post treatment verification of delivered dose using PET-imaging.

Ag-Nanogel Formulation and Performance. Ag-nanogel solutions were formulated as a homogenous dispersion composed of SAIB and EtOH (80:20) including PVP-coated AgNP (100 mg/mL). SAIB has previously been reported as an efficient and biocompatible gelation matrix for encapsulation of hydrophobic colloids[18,19] and the amount of EtOH was chosen to ensure that the Ag-nanogel formulation could be injected using thin hyperdermic needles (≤25G) to ease handling and possible patient compliance. PVP-coated AgNPs (50-80 nm, ζ-potential = 28.64±1.29 mV) were chosen as the apparent hydrophobicity of these colloids was found to be compatible with the hydrophobic SAIB based gelation matrix, a parameter previously identified as critical to prevent rapid burst release of colloids solution upon injection.[18,19]

Upon hydration, a highly viscous dark colored Ag-nanogel was formed with no apparent AgNP release despite the high colloid concentration (100 mg/mL) of PVP-coated AgNPs (Figure 1A).

In vitro release kinetics of encapsulated PVP-coated AgNPs from the formed Ag-nanogels was examined by placing small droplets of Ag-nanogels into MQ-H_2O at physiological temperature. Small aliquots of MQ-H_2O were removed over a period of 7 days and replaced with fresh MQ-H_2O to mimic the sink effect *in vivo*. AgNPs potentially undergo dissolution in aqueous environment due to oxidation by dissolved oxygen,[25,26] hence, in order to evaluate both the amount of PVP-coated AgNPs and possible Ag^+-ions formed by AgNP-dissolution inductively coupled plasma mass spectrometry (ICP-MS) was used to quantify the total amount of Ag released from the Ag-nanogels *in vitro* (Figure 1D). No detectable release of Ag^0/Ag^+ was observed confirming the high colloidal stability of the PVP-coated AgNP within the hydrophobic SAIB-based gel matrix. The Ag-nanogel

stability is an important feature, as the total amount of Ag within the Ag-nanogel is a decisive factor for subsequent dosimetric readout based on PET. Loss of Ag over time would potentially cause a drift in Ag-nanogel sensitivity and accuracy.

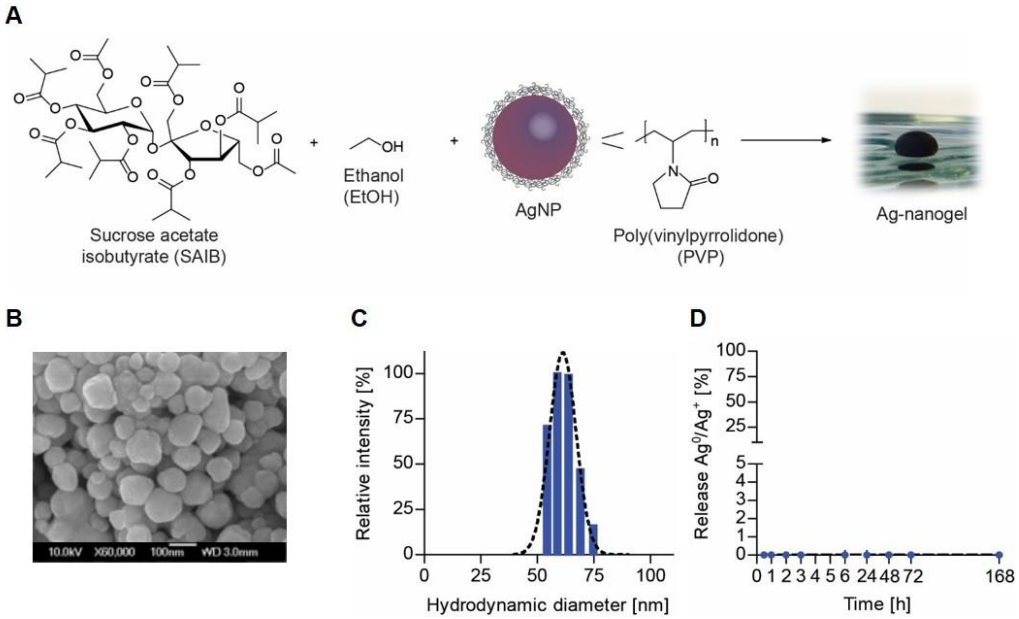


Figure 1. A) Ag-nanogel formulation based on SAIB, EtOH and PVP-coated AgNPs; B) Scanning electron microscope image of PVP-AgNP (re-printed with permission from US Research Nanomaterials Inc. Texas, USA); C) Dynamic light scattering analysis of the hydrodynamic diameter of PVP-coated AgNP and D) Total Ag (Ag^+ and Ag^0) release from Ag-nanogels *in vitro* as analyzed by ICP-MS.

***In Vitro* Evaluation of Ag-Nanogel Biocompatibility.** In order to evaluate the biocompatibility of the Ag-nanogel, cell viability of fibrosarcoma cells (HT 1080) in presence of Ag-nanogels were examined *in vitro*. Ag-nanogels and control gels based on SAIB:EtOH without PVP-coated AgNP were incubated with HT 1080 cells over a period of 24h to evaluate the Ag-nanogel influence on cell growth pattern and cell viability. Following incubation, cells were stained using the HCS Live/Dead Green Kit and fixed with 4% paraformaldehyde prior to visualization by confocal microscopy. EtOH treated cells were included as a positive control to monitor cell death (Figure 2). As evident from the *in vitro* biocompatibility assay no change in cell viability could be visualized when incubated in presence of control gel (SAIB) or Ag-nanogels. All visualized cells in close proximity to the Ag-nanogel and throughout the dish were alive and exhibited the expected cell morphology, hence, no acute toxicity indicating poor biocompatibility of the Ag-nanogel was observed. This observation is in good correlation with the undetectable release of Ag^0/Ag^+ when incubated at physiological temperature over a week.

The *in vitro* and *in vivo* biocompatibility of PVP-coated AgNP, with comparable physicochemical properties as the PVP-coated AgNPs used for Ag-nanogel preparation, has recently been thoroughly investigated by Ahlberg *et al.*[27] in a multi-center study. Adverse effects following PVP-coated AgNP exposure was predominately observed at high concentrations and in a cell-type specific manner. However, conflicting evidence in the literature regarding AgNP toxicity exist possibly due to tremendous variation in particle size, surface charge, coating methods, exposure route and exposure levels, hence, evaluation of biocompatibility must be evaluated on a case-by-case basis before clinical translation is considered feasible.[28,29,30,31]

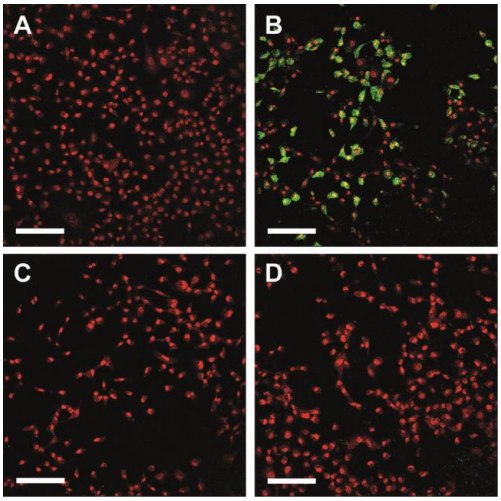


Figure 2. *In vitro* cell viability of fibrosarcoma cells (HT1080) in presence of Ag-nanogels. A) HT1080 cells only (negative control), B) Ethanol treated HT1080 cells (positive control), C) Control gel composed of SAIB and D) Ag-nanogel containing 100mg/mL PVP coated AgNP. Cells were stained with the HCS LIVE/DEAD Green Kit. Red: all cells, green: dead cells, scale bar: 150 µm.

Ag-Nanogel in External Beam Radiotherapy. Ag-nanogels with 10.0 w/w% PVP-coated AgNPs and control gels composed of the gelation matrix without PVP-coated AgNPs were casted into gelatin inside low-density polyethylene tubes. The phantom inserts were placed centrally in the mediastinum part of a tissue equivalent thoracic phantom (Figure 3A) after which the phantom was CT-scanned (Figure 3B) for subsequent radiotherapy treatment preparation and treatment planning in line with current clinical practice (Figure 3C). Ag-nanogels were easily detected on the planning CT-scan due to the high concentration of PVP-coated AgNPs within the Ag-nanogels. Ag has a high absorption coefficient of X-rays (e.g. at 100 keV; silver: 1.47 cm^2/g; bone: 0.186 cm^2/g and soft tissue: 0.169 cm^2/g),[32] which results in a high radio-opacity on the planning CT-scans. The obtained contrast level of approximately 1400-1600 HU (Table 1) is comparable to other liquid fiducial markers reported for use in image-guided radiotherapy (IGRT).[19,33] Hence, the Ag-nanogel

may in addition to its use as dosimeter, simultaneously be used in IGRT as a surrogate for tumor localization. In contrast, the control gel composed of SAIB was barely identified on the CT-scans due to low radiopacity of approximately 70 HU comparable to that of soft tissue.[34]

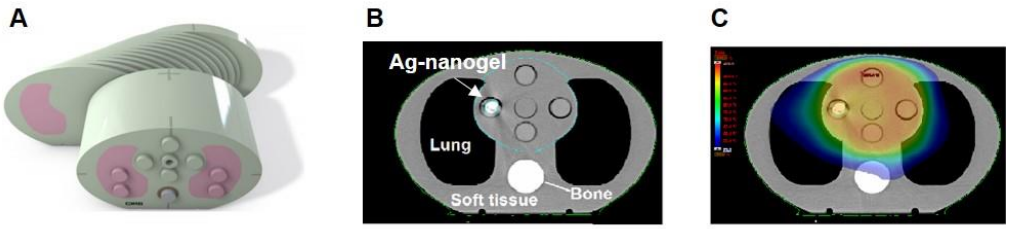


Figure 3. A) CIRS IMRT throrax phantom consisting of three tissue equivalent materials (lung (pink), soft tissue (green) and bone (white) used for Ag-nanogel dosimetry evaluation; B) Planning CT-scan of throrax phantom with Ag-nanogel inserted into the mediastinal part and visible Ag-nanogel and C) Planning target volume contoured as the entire mediastinal part of the phantom exposed to a single fraction of either 2-, 10- or 22 Gy.

Table 1. Measurements of contrast levels (HU) for each individual Ag-nanogel using 300 HU as lower threshold for automatic Ag-nanogel delineation. The CT-scans were acquired with 12-bit depth and the maximum Hounsfield Units measureable is 3071 HU.

Ag-nanogel #	Mean contrast level (HU)	STDDev	Max contrast level (HU)
1	1399	832	3071
2	1526	982	3071
3	1585	881	3071
4	1391	886	3071
Control	69	14	96
6	1397	957	3071
7	1397	952	3071
8	1519	917	3063
9	1409	793	3069
Control	69	15	103

Ag-Nanogel Dosimetric Readout using Positron Emission Tomography. During external beam radiotherapy with photon energies above 10 MV, a photonuclear reaction (γ,n) may take place where, following the absorption of the photon, a neutron is emitted from the target nucleus. Naturally occurring Ag consists of two isotopes ^{107}Ag (51.8%) and ^{109}Ag (48.2%), and therefore the (γ,n)-process may form the two radioactive Ag-isotopes; ^{106}Ag ($T_{½}$ = 23.96 min) and ^{108}Ag ($T_{½}$ = 2.37 min) (Figure 4). The activation cross-section of the two isotopes can be assumed identical[35]. The combined β^+ decay from both ^{106}Ag and ^{108}Ag can be detected in a PET scanner, however, as ^{108}Ag has a half-life of 2.37 minutes and only 0.23% of the decay from ^{108}Ag is β^+ the contributing signal from ^{108}Ag can be ignored[17]. After 14 minutes, the activity from the ^{108}Ag is less than a hundredth of that from ^{106}Ag. Additionally, as the number of formed ^{106m}Ag nuclei is negligible

compared to the total number of ^{107}Ag nuclei present within the Ag-nanogel, depletion caused by multiple irradiation incidences can be ignored.

Additional isotopes can be generated during radiotherapy namely ^{15}O ($T_{½}$ = 2.037 min) and ^{11}C ($T_{½}$ = 20.33 min), however, the photonuclear cross sections for these processes are relatively small (< 20 MeV). The influence of ^{15}O and ^{108}Ag can be minimized by allowing time for them to decay. By only using the last 20 minutes of the PET scan, we thus allow 10 minutes plus the transport time from radiotherapy to scanner for it to decay. The possible generation of ^{11}C is more problematic, as the half-life is comparable to ^{106}Ag. However, as seen from Figure 5, the control gel is not activated to any measurable degree and it would, thus, seem that a negligible amount of ^{11}C is produced.

Irradiation of the Ag-nanogel within the human equivalent thorax phantom was conducted using an approximately constant dose rate (600 monitor units/min). The dose rate is only approximate as the dose rate can change in VMAT treatments, due to occlusion by the multi-leaf collimator system. The formation of ^{106}Ag can therefor also be assumed approximately constant. The number of produced ^{106}Ag nuclei at the end of irradiation; $N(t)$ can be calculated as[17];

$$N(t) = \frac{K_F}{\lambda}\left(1 - e^{-\lambda t}\right)$$

Where K_F is the formation rate of ^{106}Ag, λ is the decay constant of ^{106}Ag, and t the time from irradiation start. As a fraction of the total number produced N_T we then get:

$$\frac{N(t)}{N_T} = \frac{N(t)}{K_F t} = \frac{K_F}{K_F \lambda t}\left(1 - e^{-\lambda t}\right) = \frac{1 - e^{-\lambda t}}{\lambda t}.$$

This fraction can be used to correct the dose delivered during external beam radiotherapy to an irradiated dose based on the amount of ^{106}Ag present at the end of irradiation.

PET-imaging of the irradiated Ag-nanogels clearly identified the Ag-nanogels due to the β^+-decay of predominately ^{106}Ag, whereas, the control gel composed of SAIB was not rendered radioactive during external radiotherapy (Figure 5C). Co-localization of the PET/CT scans showed excellent agreement of the Ag-nanogels in both the axial and sagittal plane (Figure 5B).

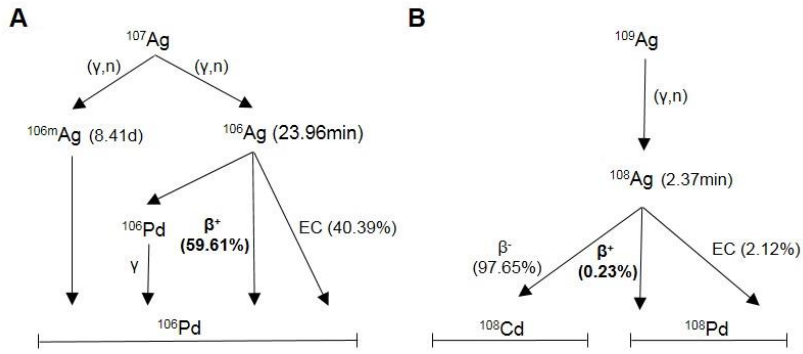


Figure 4. Nuclear transformation of naturally occurring silver; ^{107}Ag (51.8%) and ^{109}Ag (48.2%) following irradiation with a >10MV photon beam.[17]

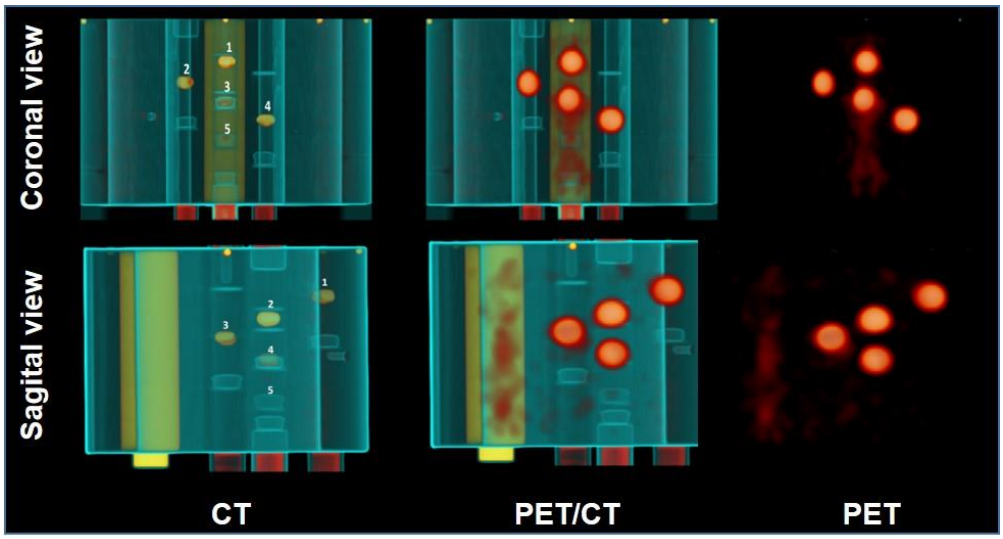


Figure 5. Coronal- and right sagittal view of 3D reconstructions of CT (left), PET/CT (middle), and PET (right) images. The gel positions are labelled on the CT reconstruction. The underlying PET images were reconstructed with a 3D iterative method for better visualization.

Statistical Analysis. With the experimental set-up, the measured activity is naturally influenced by the irradiation dose, however, potential effects caused by differences in Ag-nanogel position within the thorax phantom and variations due to inhomogeneity of the Ag-nanogel could also be present. Any differences caused by Ag-nanogel position within the thorax phantom would greatly limit the clinical potential of the Ag-nanogel as a dosimeter, as corrections would have to be applied on a patient-to-patient basis. Variations due to inhomogeneity of the Ag-nanogels could potentially be minimized by tightly controlled preparation methods.

A statistical analysis (see Supporting Information) was conducted to evaluate the potential effect of Ag-nanogel position within the phantom and any variations due to inhomogeneity of the Ag-nanogels. Initially, a general linear model was considered, where the measured activity was depending on the irradiated dose, the Ag-nanogel position within the phantom and the Ag-nanogel variations, including their pairwise interactions. Using this model, 99.32% of the total variation could be explained. However, as this model had multiple parameters, sub-models without interactions were proposed, based on; a) irradiated dose, Ag-nanogel position, Ag-nanogel variations and an intercept; b) irradiated dose and an intercept, and c) irradiated dose only (strict proportionality). These models were found to explain 99.16%, 98.25%, and 98.16% of the variability in the measured activities, i.e. only a modest loss in the ability to describe the dosimetric readout by going from the complicated to the very simple model.

In order to validate the simple model, the data set was divided into two, one containing the 9 control measurements on the same control gel in the same position within the thorax phantom and the other comprising the remaining 36 measurements recorded with varying irradiation dose, different Ag-nanogels and changing positions within the thorax phantom. As the control gel has no variability from either position or differences in gel composition, it represents an estimate of the experimental error. By comparing the estimates of the standard error for these two groups by using the F-test,[36] acceptance of the hypotheses that the "honest" error estimate based on repeated measurements of the control gel, is equal to the error estimate based on deviations from the proportional model (p-value=0.6458) was obtained. Therefore, it was concluded that the simple model just taking the irradiated dose (strict proportionality) into account was an adequate and valid description of the relationship between the measured activity (kBq/mL) and the irradiated dose (Gy);

$$activity = 0.0113 \cdot irradiated\ dose + error\ term, \qquad \hat{\sigma}_{error} = 0.01240.$$

This linear relationship between the measured activity and the irradiated dose is illustrated in Figure 6, including the 45 measurements conducted for the Ag-nanogels and the control gel. When calculating the real dose delivered to the gels from the activation using the simple proportionality model and the irradiation time, standard deviations of 0.6233 Gy (31.1 %), 1.0033 (10.0 %) and 1.7217 (7.8 %) for the 2, 10 and 22 Gy fractions regimes, respectively, are obtained. For the 2 Gy level the standard deviation is clearly too high for practical use, while it is within the acceptable range of ≤10% for 10 and 22 Gy fractions.

A possible way to reduce the standard deviation would be to increase the amount of Ag within the irradiated area either by increasing the concentration of PVP-coated AgNP within the Ag-nanogels or by increasing the volume of the Ag-nanogel. The latter could be problematic for smaller tumors with limited size. Using AgNPs composed of only ^{107}Ag would also reduce the observed standard deviation as only this isotope contributes to the signal; however, the practicality of this approach is limited.

Current commercial *in vivo* dosimetric systems are divided into two types; external- (skin) and internal (intra cavitary) dosimeters. The external *in vivo* dosimeters have been shown to have a standard deviation of up to 3%,[37] however their use in especially IMRT and VMAT is rather limited, as they require a fixed gantry position and never fully estimates the tumor dose. Intra cavitary dosimeters commonly have a standard deviation of 3.5-5%[12,37,38,39], similar to a novel colorimetric plasmonic nanosensor based on gold nanoparticles described by Pushpavanam *et al.*[40] for dosimetric use in endorectal balloons. Intra cavitary dosimeters in general do not provide information about the intra tumoral dose, as readout is only given for the cavity in which the dosimeter is placed (e.g. rectum or nasal cavity).

Compared to the commercial intra cavitary dosimeters the observed standard deviation for the developed Ag-nanogel is slightly higher, however, it has the possibility to be placed intratumorally, thereby providing valuable information about the actual delivered dose to the tumor. Additionally, the Ag-nanogel may simultaneously be used in IGRT as a tumor surrogate for motion management, and can furthermore be repeatedly used with no additional patient discomfort or added treatment time despite the PET/CT-scan required for non-invasive dosimetric readout.

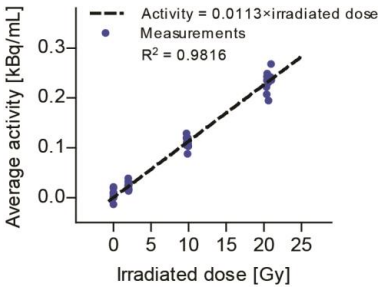


Figure 6. Correlation between irradiated dose (Gy) delivered during external radiotherapy and average measured activity (kBq/mL) recorded post-irradiation using PET-imaging.

Conclusion

A novel Ag-nanogel based on PVP-coated AgNPs formulated in a mixture of SAIB/EtOH has been developed for use as *in vivo* dosimeter for external beam radiotherapy. Irradiation of the Ag-nanogel with high-energy photon beams used clinically during treatment of cancer patients triggered the formation of ^{106}Ag and the subsequent β^+-decay could non-invasively be recorded using PET-imaging. Using a simple proportionality model, the relationship between the measured activity and the irradiated dose could be described. A high standard deviation (30.6%) was found in Ag-nanogels irradiated in the normal fractionated region (2 Gy) due to low activation of the Ag-nanogel. The low degree of ^{106}Ag formation translates into relatively higher uncertainty, which limits it's use in such treatment regimens. Irradiation of the Ag-nanogel in the stereotactic radiotherapy (10- and 22 Gy) region efficiently activated the Ag-nanogel, thereby reducing the observed standard deviation in the dosimetric readout to 10.0% and 7.8%, respectively.

In conclusion, proof-of-principle applying Ag-nanogels as in vivo dosimeters for external beam radiotherapy has been established. The developed Ag-nanogel non-invasively provided dosimetric information about the exposed dose and simultaneously provided high radio-opacity (1400-1600 HU) on the planning CT-scans sufficient for use in image-guided radiotherapy. The Ag-nanogel may provide radiation oncologist with a new tool for improving planning, precision and verification of the delivered absorbed dose and thereby provide a more precise, efficient and safer delivery of radiation to tumors.

Materials and Methods

General. PVP-coated AgNP (Ag, 99.99 %, 50-80 nm, 0.2 % w/w% PVP) was purchased from US Research Nanomaterials Inc. (Houston, TX, USA). All other chemicals were purchased from Sigma–Aldrich (Schnelldorf, Germany) and used as received from the manufacturer.

AgNP Characterization. The hydrodynamic diameter of the PVP-coated AgNPs (100 μg/mL) was analyzed using a Zeta PALS Analyzer (Brookhaven Instruments, NY, USA). Data was fitted using built-in software to calculate the hydrodynamic size of the AuNP-colloids. The ζ-potential was

measured in the same sample as the DLS analysis using a Zeta PALS Analyzer (Brookhaven Instruments, NY, USA) with a pre-conditioned electrode with 10 sub-runs while observing a fitting model residual less than 0.04.

Ag-Nanogel Formulation. SAIB (12.44 g, 90 w/w% in EtOH) and anhydrous EtOH (1.56 g, 1.98mL) were thoroughly mixed and added to PVP-coated AgNP (1400 mg) in a glass vial. The Ag-nanogel formulation was extensively vortexed and sonicated for 10 min at 50°C to form a dark colored homogeneous colloid dispersion with a final composition of SAIB/EtOH (80:20) with 10.0 w/w% PVP-coated AgNPs. The formed Ag-nanogel dispersion was stored at room temperature and used within 4h from preparation with vortexing prior to use.

In Vitro AgNP Release. Ag-nanogels were prepared as described above. Ag-nanogels (n = 3, 200µL) were added to glass vials containing MQ-H_2O (10.0 mL, 37°C) and aliquots (1.0 mL) were removed as a function of time (0.5 h, 1 h, 2 h, 3 h, 6 h, 24 h, 48 h, 72 h and 7 days) and replaced with an equal amount of MQ-H_2O. Release of Ag^0 and Ag^+ from the Ag-nanogels were measured using ICP-MS.

Ag Quantification Using ICP-MS. The Ag-concentration in aliquots removed over time from the Ag-nanogel release assay was determined by ICP-MS using a iCAPQ ICP-MS (Thermo Scientific, MA, USA) using the following settings; RF Power: 1550 W, Plasma Gas Flow: 14L/min, Nebulizer Gas Flow: 0.945 L/min, Dwell Time: 0.01s, Sweeps: 300, Replicates: 3, Delay Time: 60 s, Wash time: 60 s. Aliquots samples (50 µL) were digested for 15 min at 25°C in an open glass vial using 70% HNO_3 (1.20 mL) and diluted 25 fold with 2 % HNO_3 + 1ppb Pt as internal standard. An Ag standard curve covering the concentration range of 12.5 ppb – 100 ppb in HNO_3 (2 % HNO_3 + 1 ppb Pt) was prepared using an Ag-standard for ICP-MS (10 ppm Ag). The intensity of ^{107}Ag (m/z = 106.9050) was measured and the concentration of Ag in the aliquots determined based on the corresponding standard curve ($R^2 > 0.99$).

Ag-Nanogel Biocompatibility. MatTek 35 mm petri dishes (n = 6) with cover glass bottoms were plated with 2.0×10^4 cells/cm^2 HT1080 cells in 4 mL media and incubated overnight at 37°C. Premade Ag-nanogels (50 µL, n = 3) and SAIB based control gels (50 µL, n = 3) were added to the center of the wells. Following incubation for 24 h at 37°C the cells were stained using the HCS LIVE/DEAD Green Kit (H10290, Invitrogen, New York, USA) according to the manufactures instructions. Untreated- and ethanol treated cells (25 vol/vol% for 10 min) were included as a positive and negative control. After staining, all cells were fixed using 4 % paraformaldehyde for 15 min at 5°C, washed twice in PBS and kept in PBS for imaging using confocal microscopy (Leica TCS SP5 AOBS confocal microscope, Leica Microsystems, Wetzlar, Germany). Images were collected by simultaneously line scanning with excitation at 488- and 633 nm (green/red) and emission 493-600 nm and 638-850 nm, respectively, obtained by tunable high-reflectance mirrors.

Thorax Phantom Insert Preparation. Ag-nanogels were prepared as previously described. Ag-nanogels (1.00 mL, n = 8) giving 100 mg Ag per Ag-nanogel and control gel (1.00 mL, n = 1) (SAIB/EtOH (80:20)) were prepared by injecting the respective solutions into MQ-H_2O (20 mL) in glass vials (26 mL) at 25°C. After 5 days, a 10 w/w% solution of gelatin from bovine skin (225 G Bloom strength) was prepared by dissolving gelatin (80.0 g) and sodium azide (800 mg) in MQ-H_2O (720 mL). The mixture was heated to 80°C with stirring on a water bath until the gelatin was completely dissolved after which the solution was cooled to 40°C. Gelatin solution (10 mL) was

added to low-density polyethylene tubes (d = 24.8 mm, h = 74.5 mm) and allowed to set at 5°C for 30 min. Once set, Ag-nanogels and control gel were individually placed in each polyethylene tube and gelatin (10 mL) was added to fixate the Ag-nanogels. Thorax phantom inserts were stored at 5°C until use and in between irradiation experiments.

Thorax Phantom Setup. An IMRT Throrax Phantom (Model# 002LFC, CIRS, NJ, USA) consisting of three tissue equivalent materials (lung (pink), soft tissue (green) and bone (white)) was used for Ag-nanogel dosimetry evaluation (Figure 4A). The thorax phantom inserts, prepared as described above, were positioned in the mediastinal part of the phantom by removing the existing exchangeable cylinders. Ag-nanogels (n = 8) were randomly divided into two sets of four. For each experiment, one set was inserted into the thorax phantom. A control gel based on SAIB/EtOH (80:20) was additionally inserted into the mediastinal part of the phantom. Three irradiations were performed for each dose level (2-, 10- or 22 Gy), for a total of 36 measurements of the Ag-nanogel and 9 measurements of the control gel (see Supporting Information)

External Beam Radiotherapy Planning and Delivery. Radiotherapy treatment preparation and planning were performed in line with current clinical practice. The tissue equivalent thorax phantom with inserted Ag-nanogels was subject to a planning CT-scan using a Siemens Somatom Definition AS Scanner (Siemens, Forchheim, Germany) using a slice thickness of 2mm and tube voltage of 120 kVp. The planning target volume was contoured as the whole mediastinal part of the phantom and treatment planning was performed with volumetric-modulated arc therapy[41] with a single fraction with prescribed doses of either 2, 10 or 22 Gy. Each Ag-nanogel was contoured to retrieve contrast information of the samples and calculated dose. Each dose level was optimized individually with 18 MV photon beam using one full arc rotation (358 degrees). External beam irradiation was performed on a clinical iX Triology linear accelerator (Varian, Palo Alto, CA, USA). Prior to start of irradiation the tissue equivalent thorax phantom was positioned using image guidance to ensure a correct phantom and insert alignment.

Positron Emission Tomography Dosimetric Readout. Directly following irradiation the thorax phantom was transported to a PET/CT scanner (Biograph TrueV, Siemens, Erlangen, Germany) while carefully timing the delay time from start of irradiation to start of PET/CT-acquisition. A CT-scan was acquired for attenuation correction followed by a 30 minute PET scan in list mode. The PET-scanning was divided into three time frames; 0-2, 2-10, and 10-30 minutes and reconstructed using Filtered Back Projection (FPB) using a 4mm Hann filter and both attenuation and scatter correction. The time from irradiation end to scan start was used to correct the measured activity for decay, assuming only ^{106}Ag was present. The activity from the third frame within a sphere with a radius 20 mm centered on each gel was averaged (PMOD, PMOD Technologies Ltd., Zürich, Switzerland) and used for calculation of the dosimetric measured absorbed dose.

Statistical Analysis. Statistical analysis of the intra Ag-nanogel variance was conducted using Levene's test for equal variance[42]. Furthermore, a statistical analysis investigating the effect of Ag-nanogel position within the thorax phantom, the effect of differences in the Ag-nanogels and the delivered absorbed dose levels were conducted. All statistical analyses were performed in SAS version 9.4 (SAS Institute Inc., Raleigh, USA).

Conflicts of Interest. The author declare no competing financial interest.

Acknowledgements. This work has been funded by the Technical University of Denmark (DTU), the Niels Bohr Institute and the Danish Council for Strategic Research (Nanoguide #0603–00442B).

Supporting Information Available. The Supporting Information is available free of charge on the ACS Publications website.

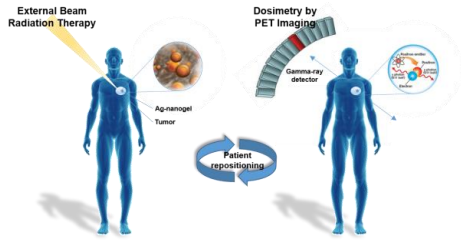
References

[1] Jemal, A.; Bray, F.; Center, M. M.; Ferlay, J.; Ward, E.; Forman, D. Global Cancer Statistics. CA *Cancer J. Clin.* **2011**, *61*, 69–90.
[2] Baskar, R.; Lee, K. A.; Yeo, R.; Yeoh, K. Cancer and Radiation Therapy: Current Advances and Future Directions. *Int. J. Med. Sci.* **2012**, *9*, 193–199.
[3] Dawson, L. A.; Sharpe, M. B. Image-guided Radiotherapy: Rationale, Benefits, and Limitations. *Lancet Oncol.* **2006**, *7*, 848–858.
[4] Bussink, J.; Kaanders, J. H. A. M.; van der Graaf, W. T. A.; Oyen, W. J. G. PET-CT for Radiotherapy Treatment Planning and Response Monitoring in Solid Tumors. *Nat. Rev. Clin. Oncol.* **2011**, *8*, 233–242.
[5] Verellen, D.; De, R. M.; Linthout, N.; Tournel, K.; Soete, G.; Storme, G. Innovations in Image-guided Radiotherapy. *Nat. Rev. Cancer* **2007**, *7*, 949–960.
[6] In *Determination of Absorbed Dose in a Patient Irradiated by Beams of X or Gamma Rays in Radiotherapy Procedures – Report 24*; International Commission on Radiation Units and Measurements (ICRU): Maryland, 1976.
[7] Kutcher, G. J.; Coia, L.; Gillin, M.; Hanson, W. F.; Leibel, S.; Morton, R. J.; Palta, J. R.; Purdy, J. A.; Reinstein, L. E.; Svensson, G. K. Comprehensive QA for Radiation Oncology: Report of AAPM Radiation Therapy Committee Task Group 40. *Med. Phys.* **1994**, *21*, 581–618.
[8] Van Dam, J.; Marinello, G. In *Methods for In Vivo Dosimetry in External Radiotherapy*; 2nd Edition; European Society for Radiotherapy & Oncology (ESTRO): Brussels, 2006.
[9] Council Directive 2013/59/Euratom. December 2003, pp. 1–73, 2014.
[10] Rikner, G.; Grusell, E. Effects of Radiation Damage on P-type Silicon Detectors. *Phys. Med. Biol.* **1983**, *28*, 1261–1267.
[11] Van Dam, J.; Leunens, G.; Dutreix, A. Correlation Between Temperature and Dose-rate Dependence of Semiconductor Response; Influence of Accumulated Dose. *Radiother. Oncol.* **1990**, *19*, 345–351.
[12] Engström, P. E.; Haraldsson, P.; Landberg, T.; Hansen, H. S.; Engelholm, S. A.; Nyström, H. In Vivo Dose Verification of IMRT Treated Head and Neck Cancer Patients. *Acta Oncol.* **2005**, *44*, 572-578.
[13] Kjær-Kristoffersen, F.; Ohlhues, L.; Medin, J.; Korreman, S. RapidArc Volumetric Modulated Therapy Planning for Prostate Cancer Patients. *Acta Oncol.* **2009**, *48*, 227-232.
[14] af Rosenschöld, P. M.; Engelholm, S.; Ohlhues, L.; Law, I.; Vogelius, I.; Engelholm, S. A. Photon and Proton Therapy Planning Comparison for Malignant Glioma Based on CT, FDG-PET, DTI-MRI and Fiber Tracking. *Acta Oncol.* **2011**, *50*, 777-783.
[15] Danos, M.; Fuller, E. G. Photonuclear Reactions. *Annual Review of Nuclear Science* **1965**, *15*, 29–66.

[16] Naseria, A.; Mesbahi, A. A Review on Photoneutrons Characteristics in Radiation Therapy with High-energy Photon Beams. *Rep. Pract. Oncol. Radiother.* **2010**, *15*, 138–144.
[17] Hansen, A. T.; Hansen, S. B.; Petersen, J. B. The Potential Application of Silver and Positron Emission Tomography for In Vivo Dosimetry During Radiotherapy. *Phys. Med. Biol.* **2008**, *53*, 353–360.
[18] Jølck, R. I.; Binderup, T.; Hansen, A. E.; Scherman, J. B.; af Rosenschöld, P. M.; Kjær, A.; Andresen, T. L. Injectable Colloidal Gold in a Sucrose Acetate Isobutyrate Gelating Matrix with Potential use in Radiation Therapy. *Adv. Healthc. Mater.* **2014**, *3*, 1680–1687.
[19] Jølck, R. I.; Rydhög, J. S.; Christensen, A. N.; Hansen, A. E.; Bruun, L. M.; Schaarup-Jensen, H.; von Wenck, A. S.; Børresen, B.; Kristensen, A. T.; Clausen, M. H.; Kjaer, A.; Conradsen, K.; Larsen, R.; af Rosenschöld, P. M.; Andresen, T. L. Injectable Colloidal Gold for use in Intrafractional 2D Image-guided Radiation Therapy. *Adv. Healthc. Mater.* **2015**, *4*, 856–863.
[20] Nischwitz, S. P.; Bauer, J.; Welzel, T.; Rief, H.; Jäkel, O.; Haberer, T.; Frey, K.; Debus, J.; Parodi, K.; Combs, S. E.; Rieken, S. Clinical Implementation and Range Evaluation of In Vivo PET Dosimetry for Particle Irradiation in Patients with Primary Glioma. *Radiother. Oncol.* **2015**, 115, 179-185.
[21] af Rosenschöld, P. M.; Capala, J.; Ceberg, C. P.; Giusti, V.; Salford, L. G.; Persson, B. R. Quality Assurance of Patient Dosimetry in Boron Neutron Capture Therapy. *Acta Oncol.* **2004**, 43, 404-411.
[22] Lu, Y.; He, H.; Cui, Y.; Tang, X. Rheological Properties of Sucrose Acetate Isobutyrate In Situ Gel. *Acta Pharma. Sinica* **2007***, 42,* 445–449.
[23] Lu, Y.; Yu, Y.; Tang, X. Sucrose Acetate Isobutyrate as an In Situ Forming System for Sustained Risperidone Release. *J. Pharm. Sci.* **2007**, *96*, 3252–3262.
[24] Lin, X.; Yang, S.; Gou, J.; Zhao, M.; Zhang, Y.; Qi, N.; He, H.; Cai, C.; Tang, X.; Guo, P. A Novel Risperidone-loaded SAIB-PLGA Mixture Matrix Depot with a Reduced Burst Release: Effects of Solvents and PLGA on Drug Release Behaviors In Vitro/In Vivo. *J. Mater. Sci.: Mater. Med.* **2012**, *23*, 443–455.
[25] Liu, J.; Wang, Z.; Liu, F. D.; Kane, A. B.; Hurt, R. H. Chemical Transformations of Nanosilver in Biological Environments. *ACS Nano* **2012**, *6*, 9887–9899.
[26] Kittler, S.; Greulich, C.; Diendorf, J.; Köller, M.; Epple, M. Toxicity of Silver Nanoparticles Increases During Storage Because of Slow Dissolution under Release of Silver Ions. *Chem. Mater.* **2010**, *22*, 4548–4554.
[27] Ahlberg, S.; Antonopulos, A.; Diendorf, J.; Dringen, R.; Epple, M.; Flöck, R.; Goedecke, W.; Graf C.; Haberl, N.; Helmlinger, J.; Herzog, F.; Heuer, F.; Hirn, S.; Johannes, C.; Kittler, S.; Köller, M.; Korn, K.; Kreyling, W. G.; Krombach, F.; Lademann, J.; *et. al.* PVP-coated, Negatively Charged Silver Nanoparticles: A Multi-center Study of Their Physicochemical Characteristics, Cell Culture and In Vivo Experiments. *Beilstein J. Nanotechnol.* **2014**, *5*, 1944–1965.
[28] Marambio-Jones, C.; Hoek, E. M. V. A Review of the Antibacterial Effects of Silver Nanomaterials and Potential Implications for Human Health and the Environment. *J. Nanopart. Res.* **2010**, *12*, 1531–1551.
[29] Wijnhoven, S. W. P.; Peijnenburg, W. J. G. M.; Herberts, C. A.; Hagens, W. I.; Oomen, A. G.; Heugens, E. H. W.; Roszek, B.; Bisschops, J.; Gosens, I.; De Meent, D. V.; Dekkers, S.; De Jong, W. H.; van Zijverden, M.; Sips, A. J. A. M.; Geertsma, R. E. Nano-silver – A Review of Available Data and Knowledge Gaps in Human and Environmental Risk Assessment. *Nanotoxicology* **2009**, *3*, 109–138.
[30] Murphy, M.; Ting, K.; Zhang, X.; Soo, C.; Zheng, Z. Current Development of Silver Nanoparticle Preparation, Investigation and Application in the Field of Medicine. *J. Nanomater.* **2015,** Article ID 696918, 12 pages.
[31] Wei, L.; Lu, J.; Xu, H.; Patel, A.; Chen, Z.; Chen, G. Silver Nanoparticles: Synthesis, Properties, and Therapeutic Applications. *Drug Discovery Today* **2015**, *20*, 595–601.

[32] Hubbell, J. H.; Seltzer, S. M. In *Tables of X-ray Mass Attenuation Coefficients and Mass Energy-Absorption Coefficients;* Version 1.4; National Institute of Standards and Technology (NIST): Maryland, 2014.
[33] Rydhög, J. S.; Jølck, R. I.; Andresen, T. L.; af Rosenschöld, P. M. Quantification and Comparison of Visibility and Image Artifacts of a New Liquid Fiducial Marker in a Lung Phantom for Image-guided Radiation Therapy. *Med. Phys.* **2015**, *42*, 2818–2826.
[34] Sanders, T. G.; Parsons, T. W. Radiographic Imaging of Musculoskeletal Neoplasia. *Cancer Control*. **2001**, *8*, 221–231.
[35] Diven, B. C.; Almy, G. M. Photo Disintegration of Silver and Aluminum. *Phys. Rev.* **1950**, *8*, 407–414.
[36] Johnson, R. A.; Miller, I.; Freund, J. In *Miller & Freund's Probability and Statistics for Engineers;* 8th Edition; Pearson, 2010.
[37] Mijnheer, B.; Beddar, S.; Izewska, J.; Reft, C. In Vivo Dosimetry in External Beam Radiotherapy. *Med. Phys.* **2013**, *40*, 070903-1–070903-19
[38] Hsi, W. C.; Fagundes, M.; Zeidan, O.; Hug, E.; Schreuder, N. Image-guided Method for TLD-based In Vivo Rectal Dose Verification with Endorectal Balloon in Proton Therapy for Prostate Cancer. *Med. Phys.* **2013**, *40*, 51715-1–51715-6.
[39] Alnaghy, S. J.; Deshpande, S.; Cutajar, D. L.; Berk, K.; Metclafe, P.; Rosenfeld, A. B. In vivo Endorectal Dosimetry of Prostate Tomotherapy using Dual MOSkin Detectors. *J. Appl. Clin. Med. Phys.* **2015**, *16*, 107-117.
[40] Pushpavanam, K.; Narayanan, E.; Chang, J.; Sapareto, S.; Rege, K. A Colorimetric Plasmonic Nanosensor for Dosimetry of Therapeutic Levels of Ionizing Radiation. *ACS Nano* **2015**, doi:10.1021/acsnano.5b05113
[41] Otto, K. Volumetric Modulated Arc Therapy: IMRT in a Single Gantry Arc. *Med. Phys.* **2008**, *35*, 310–317.
[42] Brown, M. B.; Forsythe, B. Robust Tests for Equality of Variances. *J. Am. Stat. Assoc.* **1974**, *69*, 364–367.

TOC Image


External Beam
Radiation Therapy
Ag-nanogel
Tumor
Patient
repositioning
Dosimetry by
PET Imaging
Gamma-ray
detector

Supporting Information

Injectable Silver Nanogels: *In Vivo* Dosimetry for External Beam Radiotherapy using Positron Emission Tomography

Anders Nymark Christensen[a], Jonas Scherman Rydhög[b], Rikke V. Søndergaard[c], Thomas L. Andresen[c], Søren Holm[d], Per Munck af Rosenschöld[b], Knut Conradsen[b], Rasmus Irming Jølck[c,*]

[a] DTU Compute, Department of Applied Mathematics and Computer Science, Technical University of Denmark, Building 321/324, Matematiktorvet, 2800 Kgs. Lyngby, Denmark

[b] Department of Oncology, Section of Radiotherapy, 3994, Rigshospitalet, Blegdamsvej 9, 2100 Copenhagen, Denmark and The Niels Bohr Institute, University of Copenhagen, Blegdamsvej 17, 2100 Copenhagen, Denmark

[c] DTU Nanotech, Department of Micro-and Nanotechnology, Center for Nanomedicine and Theranostics, Technical University of Denmark, Building 423, 2800 Kgs. Lyngby, Denmark

[d] Department of Clinical Physiology, Nuclear Medicine and PET, Rigshospitalet, Blegdamsvej 9, 2100 Copenhagen , Denmark

[*] Corresponding author. E-mail: rasj@nanotech.dtu.dk, Fax number: +45 45 88 77 62

1. Thorax Phantom Setup Plan

Irradiation of the thorax phantom was performed by using two different Ag-nanogel configurations where Ag-nanogel #1-4 was used in the first configuration (I) and Ag-nanogel #5-8 was used in the second configuration (II) (Table). Each dose level was irradiated and measured with the PET/CT three times. Delay between individual irradiations was minimum 24h equal to 60.1 half-lives of ^{106}Ag to prevent carry-over of radioactivity. In total 45 individual measurements were conducted, which were divided into 36 of Ag-nanogels and 9 of the control gel.

Table 1: Experimental plan with gels configuration.

Scan\Dose [Gy]	2	10	22
1	I	II	I
2	I	II	I
3	II	I	II

2. Statistical Analysis

2.1 Ag-Nanogel Variance Analysis

To analyze possible effects of the individual gels, the position in the phantom, and the irradiated dose, a model with second order interactions was considered:

$$M1: \quad Y_{ijk\nu} = \mu + pos_i + gel_j + dose_k + \{pos \cdot gel\}_{ij} + \{pos \cdot dose\}_{ik} + \{gel \cdot dose\}_{jk} + \epsilon_{ijk\nu}$$

Where $Y_{ijk\nu}$ was the measured activity of position *i*, gel *j*, dose *k* and repetition ν. $\epsilon_{ijk\nu}$ was the error and μ the overall mean. Backwards elimination was performed down to the model only including main effects:

$$M2: \quad Y_{ijk\nu} = \mu + pos_i + gel_j + dose_k + \epsilon_{ijk\nu}$$

Then a testing of the main effects was performed, down to having only the dose included, as the hypothesis was, that the activity measured was only dependent on the irradiated dose:

$$M3: \quad Y_{ijk\nu} = \mu + dose_k + \epsilon_{ijk\nu}$$

$$M4: \quad Y_{ijk\nu} = dose_k + \epsilon_{ijk\nu}$$

Normality was evaluated by inspecting the residual/quantile plots. To test for equal variance the Levene's test was used [1].

Nothing suspicious was found in the residual/quantile plots and normality was accepted. Using Levene's test for equal variance[1], the four dose levels 0-, 2-, 10- and 22 Gy, were found to have equal variance P=0.0989. Please note that the dose was not completely homogeneous within each group, so equal variance was accepted.

Four models were considered M1-M4 (see equations above), ranging from complex to very simple (Table 2).

Table 2. Statistical analysis of results for each model. DF: Degress of Freedom, RSS: Residual sum of squares, MS: Mean squared error, F: F-test value

Model	No. par.	DF	RSS	MSE	F	P(F>f)	R^2
M1	19	26	0.00247935	0.00009536			0.9932
M2	11	34	0.00297182	0.00008741	0.645	0.733	0.9916
M3	2	43	0.00643799	0.00014972	4.406	0.001	0.9825
M4	1	44	0.00676092	0.00015366	2.157	0.149	0.9816

The step from M1 to M2 was not significant. The step from M2 to M3 was significant, however the decrease in R^2 was miniscule. Taking statistical methodology very literally, one should thus accept M2. The step from M3 to M4 was also not significant. Since the decrease in variance explained were so small, we choose to accept the simplest possible model M4 as it still maintains a R^2 of 0.9816. This can be interpreted as the activity measured can be adequately explained by the effective dose alone and that position in the phantom and the individual gels has a negligible effect.

To evaluate the error the irradiated dose was calculate from the activity using the model, which was grouped into the four dose levels (Table 3)..

Table 3: Error in calculated dose

Dose group [Gy]	Standard deviation [Gy]	Maximum lower error [Gy]	Maximum upper error [Gy]
0	0.927	-1.905	1.161
2	0.612	-1.419	0.822
10	0.963	-1.543	2.189
22	1.575	-2.462	3.606

References

[1] M. B. Brown and a B. Forsythe, “Robust Tests for Equality of Variances,” *J. Am. Stat. Assoc.*, vol. 69, no. 346, pp. 364–367, 1974.

Appendix D

Effect of moderate-to-high intensity aerobic exercise on hippocampus and cortical regions in patients with mild to moderate Alzheimer's disease

Effect of moderate-to-high intensity aerobic exercise on hippocampus and cortical regions in patients with mild to moderate Alzheimer's disease

C. T. Larsen[12], K. S. Frederiksen[3], S. G. Hasselbalch[3], A. N. Christensen[2], P. Høgh[4], L. Wermuth[5], A. Lolk[5], B. B. Andersen[3], H.R. Siebner[1,6], G. Waldemar[3], E. Garde[1].

[1]DANISH RESEARCH CENTRE FOR MAGNETIC RESONANCE, COPENHAGEN UNIVERSITY HOSPITAL HVIDOVRE, DENMARK

[2] DEPARTMENT OF APPLIED MATHEMATICS AND COMPUTER SCIENCE, TECHNICAL UNIVERSITY OF DENMARK, KONGENS LYNGBY, DENMARK

[3]DANISH DEMENTIA RESEARCH CENTER, DEPT. OF NEUROLOGY, RIGSHOSPITALET, UNIVERSITY OF COPENHAGEN, DENMARK

[4]REGIONAL DEMENTIA RESEARCH CENTER, REGION ZEALAND, ROSKILDE HOSPITAL, UNIVERSITY OF COPENHAGEN, DENMARK

[5]DEMENTIA CLINIC, ODENSE UNIVERSITY HOSPITAL, ODENSE, DENMARK

[6] DEPARTMENT OF NEUROLOGY, COPENHAGEN UNIVERSITY HOSPITAL BISPEBJERG, COPENHAGEN, DENMARK

Corresponding author

Christian T. Larsen

Danish Research Center for Magnetic Resonance

Section 714

Copenhagen University Hospital

Hvidovre

Kettegaard Allé 30

2650 Hvidovre

Denmark

cthla@dtu.dk

Abstract

Background: Studies on healthy elderly have shown that aerobic exercise has a positive effect on both brain structure and function. So far studies in patients with Alzheimer's disease (AD) are few and results have been inconsistent. In this study, we wanted to assess the relationship between aerobic exercise, brain changes measured by MRI and cognitive functioning in patients with AD.

Methods: As part of a larger randomized controlled trial this MR-sub-study included forty-two patients. For both control and exercise group MR and cognitive assessment was performed at baseline and after 16 weeks with 60-minutes exercise sessions three times a week. Both attendance and intensity were monitored providing a total exercise load. Changes in regional brain volumes and cortical thickness were analysed using Freesurfer and volume of white matter hyperintensities (WMH) quantified.

Results: Exercise load showed a positive correlation with changes in volume in the hippocampal subfields, as well as frontal, cortical thickness in the exercise group. Changes in frontal, cortical thickness correlated with measures of mental speed and attention (SDMT) and verbal fluency in both groups. Volume of WMH were associates with changes in hippocampal volume.

Conclusion: In patients with AD the effect of exercise on hippocampal volume appear to depend on training attendance and intensity. The extent of WMH may modify the effect of physical training but further studies are needed.

Introduction

Alzheimer's disease (AD) is a neuro-degenerative disease, characterized by progressive impairment of memory [GM84] and atrophy of specific brain regions, in particular the hippocampus [JB07]. Atrophy of the anterior hippocampus can be observed in patients with mild cognitive impairment (MCI) as early as three years prior to onset of AD, with increasing involvement of the hippocampus as the symptoms progresses [JW07].

In addition, atrophy has been observed in the amygdala, entorhinal cortex and fusiform gyrus in MCI, progressing to the middle temporal gyrus, posterior temporal lope and parietal lobe in patients with AD [JW07].

The effect of current pharmacological treatments of AD are at best symptomatic [RC12] but recent studies suggest that non-pharmacological approaches such as physical exercise may have a beneficial effect on cognitive functioning as well as brain structure.

In healthy elderly both cognition, physical functioning and performance in activities of daily living were improved when given a home training program consisting of daily exercises and walking [AV12]. Studies including MRI suggest that physical training is associated with increased whole brain volume [SC06], less atrophy in frontal, parietal and temporal cortex [SC03] and even an increase in grey matter volume in pre-frontal and cingulate cortex [RR11]. In preadolescent children Chaddock et al [LC10] showed a relation between basal ganglia volume, enhanced cognitive functioning and aerobic fitness while in healthy elderly improvement in memory function were associated with increase hippocampal volume [KE11].

In patients with AD Andrade et al reports an increase in frontal cognitive function, after following a multimodal exercise program for 16 weeks [LA13].

Increasing evidence support that exercise benefits brain function and structure but also that there are multiple pathways and that age and concurrent pathological processes may modify the effect. In a recently published study exercise had a positive effect on neuropsychiatric symptoms and cognition in a relatively large group of patients with mild to moderate (KH 2016). To our knowledge, this is the first to investigate the effect of continuously supervised moderate to high-intensity exercise program in patients with mild to moderate AD. In a sub-study, MRI was performed at baseline and 16-week follow-up in order to assess the effect on regional brain volumes. The primary outcome measure is hippocampal volume and we hypothesize that in AD patients a moderate to intense exercise program will preserve hippocampal volume.

Methods

Participants and study design

The primary objective of the ADEX study was to assess the effect of moderate-to-high intensity aerobic exercise on cognitive and physical functioning, quality of life and ADL in two-hundred community-dwelling patients with mild to moderate AD. The participants were randomized into a control and exercise group, the latter performing 60 minutes of moderate-to-high-intensity aerobic exercise three times weekly for sixteen weeks. Psychological, cognitive and physical performance was assessed before and after the sixteen week period for both groups. The trial ran for 5 rounds from 2012 to 2014.

The procedure for screening, as well as inclusion and exclusion criteria, is described in a recent publication [KH13]. In brief, key inclusion criteria included age between 50 and 90 years and a Minimal Mental State Examination (MMSE) score of more than 19, whereas exclusion criteria included presence of medical and psychiatric diseases, alcohol abuse and regular, weekly high-intensity exercise.

A subgroup consisting of seventy-one patients from memory clinics in Copenhagen, Roskilde and Odense was invited for brain MRI at baseline and 16 week follow-up. Thirteen of these patients left the study prematurely, and sixteen patients were excluded due to poor MRI data quality (movement artefacts) (9), data processing problems (6) or notable error in data processing outcomes (1), leaving 42 patients for the present study.

The ADEX trial was approved by the The Committees of Biomedical Research Ethics for the Capital Region (Protokol no.: H-3-2011-128) and by the Danish Data Protection Agency (j.no.: 30-0718).

MRI acquisition

Both baseline and follow-up MRI was performed at Hvidovre Hospital, Denmark, using a 3.0-T Siemens Tim Trio scanner and included T1-weighed magnetization-prepared rapid gradient echo (MPRAGE) (TE 3.04ms, TR 1550ms, FoV read 256mm, FoV phase 100%, 192 slices), T2-weighed fast spin echo (TE 354ms, TR 3000ms, FoV read 282mm, FoV phase 76.6%, 192 slices) and fluid attenuated inversion recovery (FLAIR) (TE 353ms, TR 6000ms, FoV read 282mm, FoV phase 85.9%, 192 slices) sequences.

Data processing

Regional, cortical thickness and hippocampal volume

The T1-weighed data was gradient unwarped to correct for spatial distortions [JC06], and then processed with version 5.3 of the cross-sectional [BF02] and longitudinal [MR12] Freesurfer stream, in order to obtain segmentations of cortical regions defined according to the Destrieux atlas [CD10] as well as the hippocampal subfields [KL09], caudate and putamen. The pipeline was specifically tuned to correct for intensity inhomogeneity that can be observed at 3T [RB08, WZ09].

In cases where Freesurfer failed to properly delineate the white matter and pial surface, the pipeline were manually guided following the steps outline in the Freesurfer documentation (http://freesurfer.net/fswiki/FreeSurferWiki). This specifically involved correcting the skull stripping to better delineate the pial surface, insertion of control points to guide white matter normalization for the purpose of improving white matter segmentation, and finally editing the white matter segmentation itself. Two trained readers edited the pipeline; to avoid segmentation bias, one was responsible for skull stripping and white matter editing, while the other was responsible for control point insertion.

Finally, overall quality of the longitudinal segmentation output were asserted by experienced raters (CTL, KSF, EG). Specifically, the pial and white matter surface outlines, as well as the hippocampal subcortical segmentation were visually inspected and consensus reached for all. One volume was excluded due to significant segmentation error in the hippocampus.

To explore regional, cortical effects, gyri and sulci thickness measures obtained from Freesurfer were divided into four categories (early, middle, late, and very late) each including areas reported to be progressively affected by atrophy from mild cognitive impairment MCI to full AD diagnosis [JW09]: 'early' (temporal, precuneus, cingulate), 'middle' (parietal, temporal-occipital, occipital, fusiform, parahippocampus) and 'late' (frontal). A 'very late' region composed by the pre and postcentral cortex were also defined (supplementary material, table 4).

Whole and parenchymal brain volume

Freesurfer also provides measures of brain volume (BV), brain parenchymal volume (BPV), white matter volume (WM) and intracranial volume (ICV). Whole brain volume included all segmented structures, excluding background and the brain stem. Parenchymal volume further excluded the ventricles (lateral, inferior lateral, 3rd, 4th and 5th), CSF and choroid plexus. Brain parenchymal fraction (BPF) was obtained by dividing BPV with ICV.

White matter hyperintensities

For delineation of white matter hyperintensities (WMH), MPRAGE and T2-weighted images were co-registered and re-sliced to the corresponding FLAIR image using a 6 parameter rigid transformation. WMH were defined as clearly hyperintense areas relative to surrounding white matter on both FLAIR and T2-weighted images and identified by simultaneous inspection of both aligned images. For WMH volume local thresholding was applied and WMH volumes for the whole brain quantified automatically using the Jim image analysis package, Version 6.0, (Xinapse Systems Ltd., Northants, UK, www.xinapse.com). Visual identification and delineation was carried out by a single trained rater blinded to clinical information. For nine subjects (five control, four intervention) WMH could not delineated due to movement artefacts.

Longitudinal and normalized measures

Longitudinal measures of brain volume, cortical thickness and cognitive scores for each subject were computed as the *relative* change between baseline and follow-up by subtracting baseline from follow-up, and dividing the difference with the baseline measure, thereby canceling out within-subject correlations, as well as accounting for between-subject differences in brain size. Throughout the paper, we will refer to the relative change simply as *change*.

A normalized WMH measure was obtained by dividing WMH volume with white matter volume.

Cognitive outcome measures

Cognitive assessment included the Minimal Mental State Examination (MMSE) for global cognitive impairment [MF75], the Symbol Digit Modalities Test (SDMT) for mental speed and attention [AS82] (only measurements at 120 seconds included in the analysis), and the Stroop Color and Word Test (Stroop) incongruent score for reaction time. Verbal memory performance was assessed by the Alzheimer's Disease Assessment Scale – Cognitive Subscale (ADAS-Cog) [WR84], and verbal fluency (VFT) as number of words produced over 1 minute each ([KH15] for details).

Exercise load (attendance and intensity)

To assess training attendance and intensity a training log was created. Attendance was logged, and attendance ratio defined as number of attended exercise sessions over total number of offered sessions. Exercise intensity was based on the per-session average heart rate (HR) recorded using continued monitoring during exercise (including rest). Average HR for all sessions was calculated, and intensity defined as average HR over maximum expected HR (220 minus subject age). To obtain total exercise load, measures for attendance ratio and intensity was multiplied.

Statistical analysis

Brain volume measures

Separate multivariate models were used to compare changes in volume between groups for the hippocampal subfields (model 1), para-hippocampus (model 2), caudate and putamen (model 3). Similarly, separate models were used to compare changes in thickness of the cortical gyri and sulci respectively for each of the 'early', 'middle', 'late' and 'very late' categories previously described.

For all group tests, Hotellings T^2 multivariate test [HH31] was applied. Hotellings T^2 eliminates the need for testing each individual measure in a model (e.g., each of the hippocampal subfields), and consequently the need for performing a multiple comparisons test. This makes the test more sensitive and less prone to type II error (false negatives) than e.g., univariate tests with Bonferoni correction.

Since outliers were detected in scatter plots of the variables, a further non-parametric Oja rank test [HO04] were performed, using 10.000 permutations, to confirm validity of p-values from Hotellings T^2 test.

Correlation tests

The relationship between changes in hippocampal subfield volume and verbal memory (ADAS-Cog measure) was tested with caudate and putamen as a control regions. Also, the relationship between changes in frontal and cingulate cortical thickness, and verbal memory, mental speed and attention (SDMT, VFT, Stroop) were assessed with pre- and postcentral cortex as a control region. The relationship between cognitive measures and changes in volume of caudate and putamen were explored with hippocampus as a control volume. Finally, relationship between exercise load and changes in hippocampal subfield volume as well as frontal cortical thickness were also investigated for the exercise group only. (Details on cortical labels can be seen in the supplementary material, table 5).

All correlation tests were performed by computing the covariance matrix between changes in test scores and brain measures, and then testing the nul-hypothesis of zero covariance between the two. This yields an overall p-value for the full covariance matrix, but not an r-value. If the overall p-value was significant, a post-hoc analysis of the correlation for each single brain measure was performed, yielding individual r- and p-values.

Six subjects were excluded from the correlation tests due to missing cognitive scores (one subject lADAS-Cog; three subjects SDMT; one subject VFT; six subjects Stroop).

For all tests, the significance level was 0.05. Gender, age and baseline WMH were used as covariates. Statistics were obtained with SAS Statistical Software version 9.4 and Rstudio 2.15.2

Results

No significant differences were found for any baseline characteristics between the control and

intervention group (Table 1).

Brain volumes

In the hippocampal subfield model a significant difference were found for the left fimbria (p=0.012) and CA2_3 (p=0.016) which however, could not be found when correcting for multiple comparisons (figure 1, table2). No difference between groups was observed for the parahippocampal or caudate and putamen models (table 2). No significant between-group difference in changes in regional cortical thickness was found (table 3).

The normalized WMH measure did not change significantly from baseline to follow-up (p=0.996). In both groups WMH was associated with changes in the hippocampal subfield volume (p=0.002), specifically presubiculum (r=-0.345, p=0.031), and CA4_DG (r=-0.406, p=0.010), (supplement, table 6) as well as changes in both gyri (p=0.048) and sulci (p=0.002) thickness in the 'very late' category, and also sulci thickness in the 'middle' category (p=0.0495). Inspection of the individual significant gyri and sulci showed the largest correlations with the right post-central gyri (r=0.221, p=0.170) in the 'very late' gyri, the right post-central sulci (r=0.467, p=0.002) in the 'very late' sulci category, and the right occipital superior and transversal sulci (r=0.460, p=0.003) in the 'middle' category.

Cognitive performance correlations

Change in the frontal and cingulate cortical thickness correlated significantly with both SDMT (p=0.025) and VFT (p=0.026), (table 7, supplement). Specifically, for SDMT a moderate correlation was found with change in cortical thickness of the right frontal inferior-orbital gyri (r=0.464, p=0.004) and right frontal inferior-triangular gyri (r=0.386, p=0.020). Per-group investigation (figure 2) revealed moderate correlations in both regions for the exercise group but not the control group. For VFT, correlations was found with cortical thickness changes in the left (r=0.384, p=0.017) and right (r=0.328, p=0.044) frontal mid-posterior gyri and sulci. Per-group investigation of this relationship (figure 3) revealed significant correlations in both regions for the control group but not the exercise group. (table 8, supplement).

A separate analysis of the correlation between normalized WMH and each cognitive performance measure showed a significant correlation with Stroop (r=0.394, p=0.023). Inspection of the scatterplot of the measures revealed the correlation to be dominated by two points with no apparent overall trend. No significant correlations were found between WMH and the SDMT/VFT measures.

Exercise load correlations

Exercise load was found to associate significantly with changes in cortical thickness in the frontal cortex (p=0.0106), especially for the right frontal inferior sulci (r=0.514, p=0.034). Similarly, a significant correlation with changes in volume in the hippocampal subfields (p=0.0091) was found, the strongest correlation showing in the right subiculum (r=0.443, p=0.086).

Discussion

To our knowledge, this is the first study to explore the effects of supervised moderate-to-high intensity aerobic exercise on regional brain atrophy measures in patients with mild to moderate AD.

Main findings in this study are three-fold. First, exercise load shows a positive correlation with changes in volume in the hippocampal subfields, as well as frontal, cortical thickness which support that exercise does stimulate brain growth, which agrees with previous findings, e.g., [SC03], [SC06] and [KE11]. The group differences in changes in the left fimbria and CA2_3 hippocampal subfield volume

disagree (figure 1). While changes in the left fimbria could suggest that exercise stimulates brain growth with measurable effects already after 16 weeks in patients with AD, the opposite trend in CA2_3 suggest that the effects are spurious.

We performed a post-hoc qualitative inspection of changes in hippocampal, caudate and putamen volume and regional cortical thickness. Our observations suggested only a minor loss of tissue across subjects across a four month period, and furthermore that loss of tissue seems to be less in exercise participants. In some cases, the data suggested a slight increase. However, given the variation in data and the inability to show significant differences between groups, this remains only partially indicative of the effects of exercise.

Previous literature, e.g., Erickson et al, who showed an effect in global hippocampal volume in healthy elderly after 1-2 years of PE [KE11]. Our findings suggests that 16 weeks may not be enough to effect duration as well as intensity should be considered when planning exercise programs. An alternative explanation for the disagreement is that the control group may have been exercising outside of the study, thereby diminishing group differences.

Second, for all participants, changes in frontal, cortical thickness were associated with SDMT and VFT. Cortical thinning have previously been shown to associate with cognitive impairment, e.g., in Parkinson's Disease [BS14], which suggests that a decline in cortical thickness can be used as an indicator of progressive cognitive impairment. Our finding that frontal, cortical thickness associates with SDMT/VFT cognitive performance measures, is in agreement with this, and more generally with literature on the role the frontal cortex has in mental speed, attention and verbal fluency [JA06].

Given that AD is a neuro-degenerative disease, it would be expected that changes in cortical thickness and cognitive measures would be primarily negative. However, no particular decline or increase in cortical thickness or cognitive performance were observed for changes in frontal, cortical thickness and the SDMT/VFT scores (figure 2 and 3). Further inspection reveals a stronger relationship between changes in cortical thickness and SDMT for the exercise group, while the control group exhibits the strongest correlation with VFT.

Third, interestingly, normalized WMH values were associated with changes in hippocampal volume as well as regional, cortical thickness in both groups, suggesting that pathology other then AD may influence brain structure. WMH are generally regarded a marker of small vessel disease and severe WMH an indicator of poor vascular health. Recent findings (presented at AAIC 2015) suggest that the presence of WMH may modify the effect of exercise intervention but also that exercise may enhance vascular health as well as connectivity.

A recent study [EL15] have shown that the rate of change in WMH are strongest during conversion from MCI to AD diagnosis, and follows the rate of change in hippocampal volume suggesting that WMH may play a part in the conversion process. Similarly, Caso et al [CF15] conclude that WM degeneration may be an early marker of pathological changes in atypical AD. In line with these studies, our finding that the severity of WMH associates with relative changes in the hippocampal subfields implies a relationship between WMH and the rate of neuro-degeneration in AD.

WMH has also been shown to associate with cognitive decline [OC10] and more specifically medial temporal atrophy, attention and frontal executive functions [YS11] as well as frontal atrophy and reduced delayed recall performance [IM12]. Here, we observed positive correlations between WMH and changes in cortical thickness, which is not what we would expect. The finding was made predominantly in the cortical control regions, where effects due to AD would be expected to show very late. As such, the finding could be spurious, and may again suggest that a four month period is not be

enough to measure significant differences with controls.

Although the underlying mechanism is still not completely understood animal studies indicate that physical exercise stimulates neurogenesis and formation of brain-derived neuro-protective factor (BDNF) [HP05][KN09][MM13]. This relationship between exercise and BDNF has also been reported in humans [CC02].

Methodological considerations

The disagreement between findings in the left fimbria and left CA2_3 hippocampal subfields, the positive correlations between WMH and regional, cortical thickness suggest some level of uncertainty in the statistics, one explanation being noise. Furthermore, the population size after drop-outs, data processing and quality assurance in this study was quite small, which affects statistical power and consequently increases the likelihood of spurious results.

From a clinical perspective, sixteen weeks of exercise are likely not enough to measure effects on the same level as e.g., Erickson et al [KE11]. Furthermore, previous findings have been in healthy elderly subjects, with a presumably small pathological load. It is very likely the effects of an exercise intervention on the brain might be different in patients with AD.

The subjects in this study were recruited from three different memory clinics. This could pose a potential bias in statistics, given that subjects may be treated differently at each center. We did not control for this, due to the already mentioned statistical concerns. Furthermore, there is a good agreement between the statistics of the control and intervention group as shown in Table 1. Given that subjects from all centers using a controlled, randomized process, it seems reasonable to assume that differences between subjects due to memory centers even out.

As described in [MR12], the longitudinal Freesurfer pipeline utilizes cross-sectionally processed time points to generate a common template, which is then used as a point of initialization for an unbiased analysis of each individual timepoint. This procedure helps to avoid potential bias in the outcome measures due to e.g., registration to the baseline time point, as pointed out in [NF11]. Furthermore, it increases statistical power because inter-subject variation is reduced.

Conclusion

A sixteen week intervention of moderate-to-high aerobic exercise clearly translated to observable changes in hippocampal subfield volume and cortical thickness in a group of patients with AD. The finding supports evidence that exercise stimulates exercise, which may potentially have a negating effect on neuro-degeneration.

Furthermore, correlations between changes in the frontal cortex and mental speed and verbal fluency show that changes in brain volume and cortical thickness does relate to changes in cognition. As such, it seems likely that exercise for prolonged periods of time may increase the extent to which brain growth is stimulated, thereby leading to a positive effect on cognition and ADL in patients with AD.

Finally, an observed association between the extent of WMH and changes in the hippocampus suggest that WMH may be indicative of the rate of neuro-degeneration. It is further suggested that this may have a limiting effect on the effectiveness of exercise.

Future studies should explore the effect of aerobic exercise focusing on a prolonged duration of the intervention period, as well as an increase in the number of participants receiving MRI, in order to increase statistical power.

Acknowledgements

The ADEX study is supported by a grant from The Danish Council for Strategic Research (j. no.: 10-092814).

Investigator list by site:

Memory Clinic, Copenhagen University Hospital, Rigshospitalet

Birgitte Bo Andersen, DMSc, M.D.

Memory Clinic, Roskilde Hospital

Peter Høgh, PhD, M.D.

Dementia Clinic, Odense University Hospital

Anette Lolk, Assoc. Professor, PhD, M.D.

Lene Wermuth, M.D.

Department of Geriatrics, Svendborg Hospital

Søren Jakobsen, M.D.

Department of Geriatrics, Slagelse Hospital

Lars P. Laugesen, M.D.

Robert Graff Gergelyffy, M.D.

Memory Clinic, Aarhus University Hospital

Hans Brændgaard, M.D.

Hanne Gottrup, PhD, M.D.

Memory Clinic, Aalborg Hospital

Karsten Vestergaard, M.D.

Memory Clinic, Glostrup University Hospital

Eva Bjerregaard, M.D.

We are grateful to all physiotherapists, study nurses and clinical raters for their contributions to the study. We also thank Jonathan Polimeni from the Athinoula Martinos Center, Massachussetts General Hospital, Boston, USA for supplying gradient unwarping software.

Figures and tables

Table 1: Baseline demographics for all participants and basic volumetric measures for participants in the MR sub-study.

	MR-substudy			Main study	
	Control (N=20)	**Exercise (N=22)**	**p-values**	**Control (N=93)**	**Exercise (N=107)**
Age (years), *mean (±SD)*	69 (7.5)	68 (7.7)	0.681	71.3 (7.3)	69.8 (7.4)
Gender, *male/female (N)*	12 / 8	14 / 8	0.809	57/36	56/51
MMSE, *median (±SD)*	26.0 (2.3)	26.0 (3.4)	0.712	24.1 (3.8)	23.8 (3.4)
Hypertension*, *N (%)*	5 (25)	4 (18)	0.591	35 (37.6)	48 (44.9)
WMH ($mm^3 \cdot 10^3$), *median (±SD)*	1.94 (6.1)	1.16 (6.3)	0.116	N/A	N/A
WMH/WM (10^{-3}), *median (±SD)*	6.00 (15.8)	2.80 (12.2)	0.085	N/A	N/A
BV ($mm^3 \cdot 10^6$), *mean (±SD)*	1.01 (0.13)	1.11 (0.12)	0.489	N/A	N/A
BPV ($mm^3 \cdot 10^6$), *mean (±SD)*	1.00 (0.11)	1.05 (0.11)	0.226	N/A	N/A
BPF, *mean (±SD)*	0.63 (0.03)	0.64 (0.03)	0.456	N/A	N/A

* Hypertension was defined as ≥140/90 mmHg.

WMH: White matter hyperintensities, WM: white matter, BV: brain volume, BPV:brain parenchymal volume, BPF: brain parenchymal fraction.

Table 2: Difference between control and exercise group in change in volume measures.

		Left side (p-value)	Right side (p-value)
Test 1	Hippocampus		
	Total hippocampus volume *	0.930	0.266
	Presubiculum	0.747	0.817
	Subiculum	0.485	0.469
	Fimbria	0.012	0.423
	Hippocampal fissure	0.240	0.583
	CA1	0.470	0.237
	CA2_3	0.016	0.299
	CA4_DG	0.120	0.184
Test 2			
	Parahippocampus	0.462	0.962
Test 3			
	Caudate	0.128	0.071
	Putamen	0.289	0.779

* Includes the 'hippocampus' class, which captures voxels that were not put in the other subfield categories.

Table 3: Difference between control and exercise group in change in cortical, regional thickness.

	Gyri (p-value)	Sulci (p-value)
Early	0.255	0.187
Middle	0.112	0.623
Late	0.687	0.126
Very late	0.121	0.506

Figure 1: Boxplots showing the changes in volume in the left fimbria and left CA2_3 hippocampal subfields for the control and exercise groups.

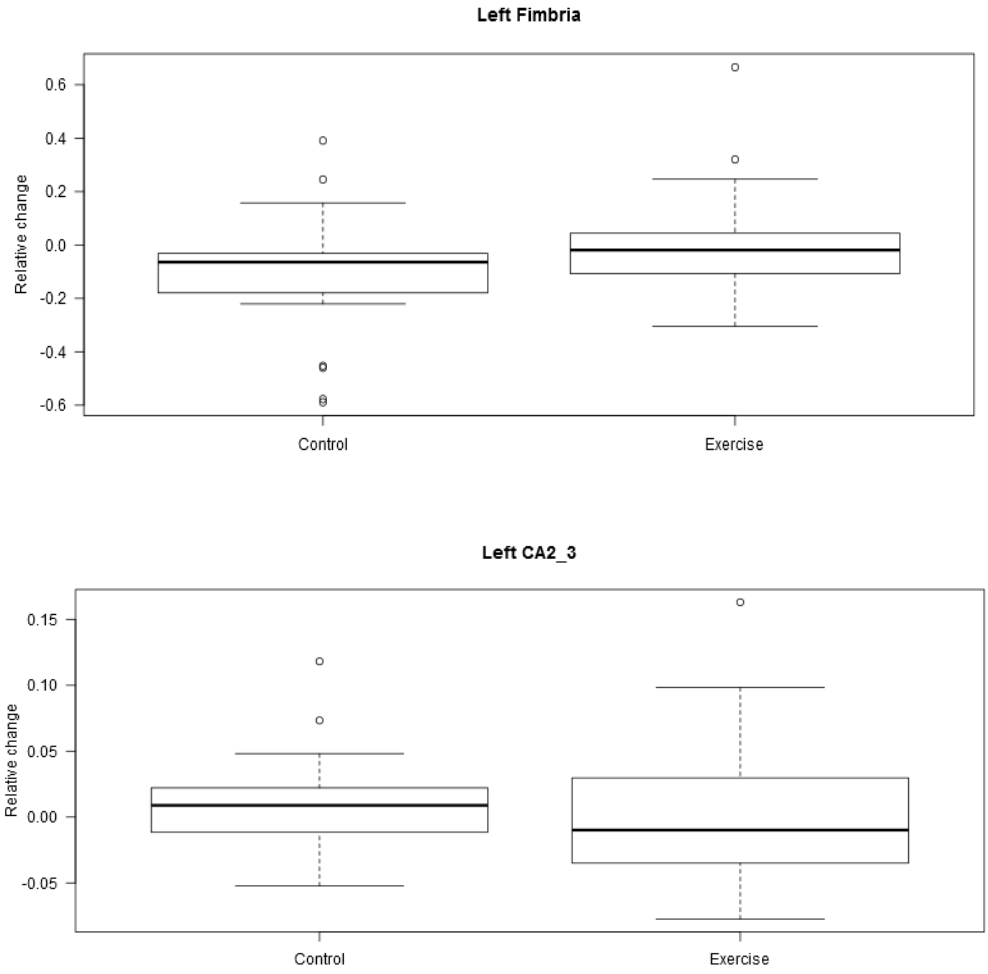


Figure 2: Correlation between change in the right frontal inferior-orbital/inferior-triangular gyri thickness and relative change in SDMT measured at 120 seconds.

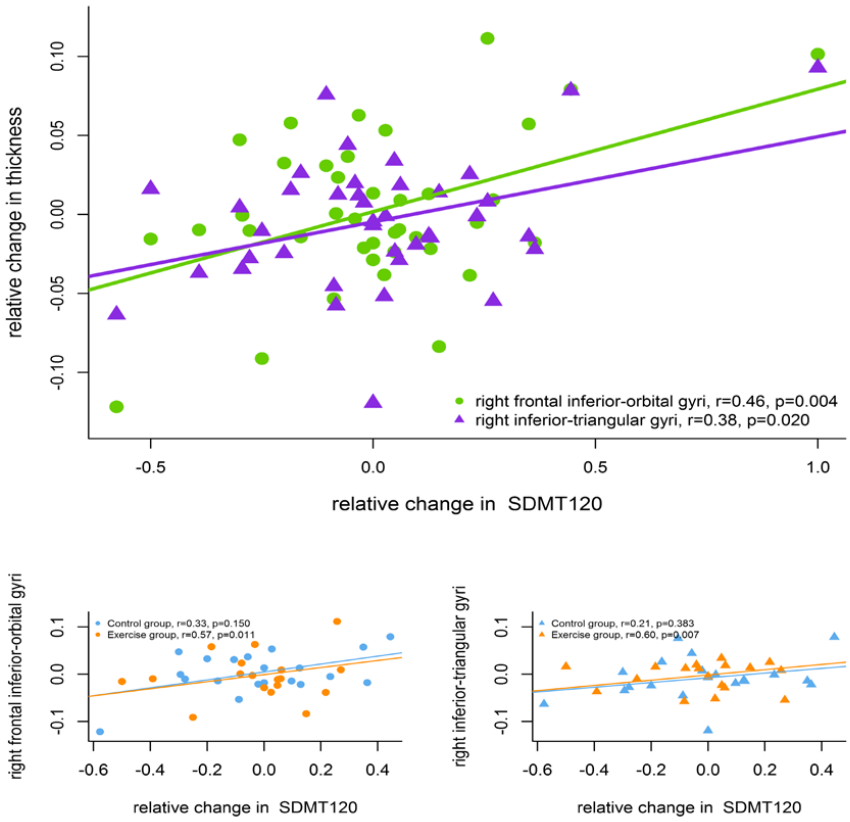
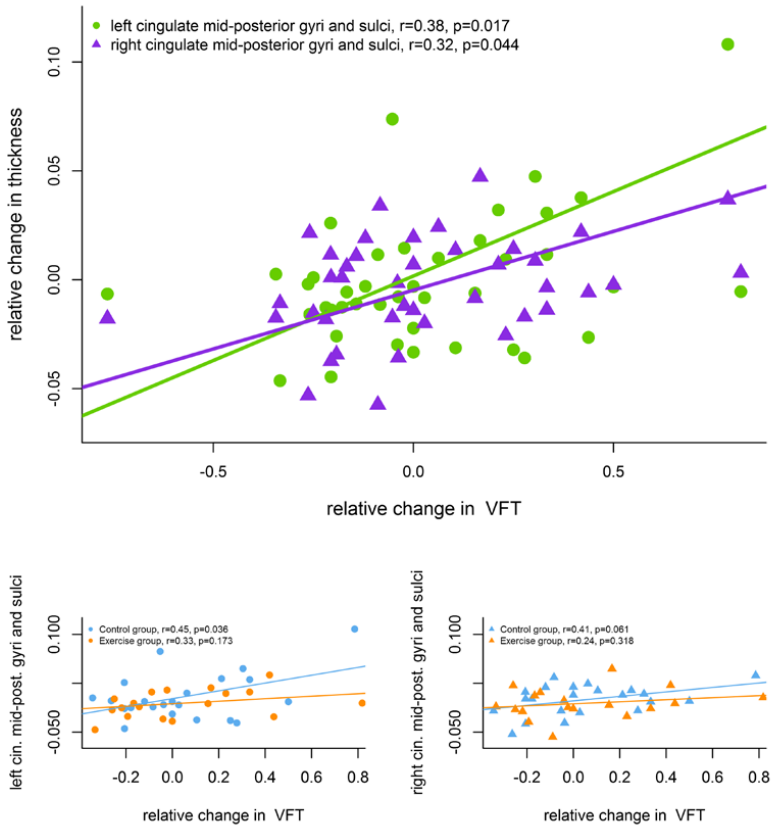


Figure 3: Correlation between relative change in the left and right cingulate, mid-posterior gyri and sulci thickness and relative change in VFT.



References

[JA06] Alvarez JA, Emory E: Executive function and the frontal lobes: A meta-analytic review. Neuropsychology Review 2006;16(1);17–42.

[LA13] Andrade LP de, Gobbi LTB, Coelho FGM, Christofoletti G, Costa JLR, Stella F: Benefits of Multimodal Exercise Intervention for Postural Control and Frontal Cognitive Functions in Individuals with Alzheimer's Disease: A Controlled Trial. J Am Geriatr Soc. 2013;61(11):1919-26.

[JB07] Barnes J, Godbolt AK, Frost C, Boyes RG, Jones BF, Scahill RI, Rossor MN, Fox NC: Atrophy rates of the cingulate gyrus and hippocampus in AD and FTLD. Neurobiol Aging 2007;28(1):20-28.

[CC02] Cotman CW, Berchtold NC. Exercise: a behavioral intervention to enhance brain health and plasticity. *Trends Neurosci*. 2002 Jun;25(6):295-301 .

[RC12] Castellani RJ, Perry G: Pathogenesis and disease-modifying therapy in Alzheimer's disease: the flat line of progress. Arch Med Res 2012;43: 694–698.

[FC15] Caso F, Agosta F, Mattavelli D, Migliaccio R, Canu E, Magnani G, Marcone A, Copetti M, Falautano M, Comi G, Falini A, Filippi M: White Matter Degeneration in Atypical Alzheimer Disease. Neuroradiology 2015.

[LC10] Chaddock L, Erickson KI, Prakash RS, VanPatter M, Voss MW, Pontifex MB, Raine LB, Hillman CH, Kramer AF: Basal Ganglia Volume is associated with Aerobic Fitness in Preadolescent Children. Developmental Neuroscience 2010;32;249-256.

[OC10] Carmichael O, Schwarz C, Drucker D, Fletcher E, Harvey D, Beckett L, Jack CR Jr, Weiner M, DeCarli C: Longitudinal changes in white matter disease and cognition in the first year of the Alzheimer disease. Arch Neurol. 2010;Nov;67(11):1370-1378.

[SC03] Colcombe SJ, Erickson KI, Raz N, Webb AG, Cohen NJ, McAuley E, Kramer AF: Aerobic Fitness Reduces Brain Tissue Loss in Aging Humans. J Gerontol A Biol Sci Med Sci. 2003;58(2):176-80.

[SC06] Colcombe SJ, Erickson KI, Scalf PE, Kim JS, Prakash R, McAuley E, et al: Aerobic exercise training increases brain volume in aging humans. J Gerontol A Biol Sci Med Sci 2006;61:1166–1170.

[JC06] Jovicich J, Czanner S, Greve D, Haley E, Kouwe A, Gollub R, Kennedy D, Schmitt F, Brown G, MacFall J, Fischl B, Dale A: Reliability in Multi-Site Structural MRI Studies: Effects of Gradient Non-linearity Correction on Phantom and Human Data. NeuroImage 2006;30(2):436-43.

[CD10] Destrieux C, Fischl B, Dale A, Halgren E: Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. NeuroImage 2010;53(1);1-15.

[KE11] Erickson KI, Voss MW, Prakash RS, Basak C, Szabo A, Chaddock L, et al: Exercise training increases size of hippocampus and improves memory. Proc Natl Acad Sci USA 2011;108:3017–3022.

[BF02] Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, van der Kouwe A, Killiany R, Kennedy D, Klaveness S, Montillo A, Makris N, Rosen B, Dale AM: Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron 2002;33:341-355.

[NF11] Fox NC, Ridgway GR, Schott JM: Algorithms, atrophy and Alzheimer's disease: Cautionary tales for clinical trials. NeuroImage 2011;51:15-18.

[MF75] Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12(3);189-198.

[CG78] Golden C. Stroop color an word test manual. Chicago; Stoelting Co; 1978.

[HH31] Hotelling H. The Generalization of Student's Ratio. Ann. Math. Statist. 2 1931;3;360-378.

[KH13] Hoffmann K, Frederiksen KS, Sobol NA, Beyer N, Vogel A, Simonsen AH, Johannsen P, Lolk A, Terkelsen O, Cotman CW, Hasselbalch SG, Waldemar G: Preserving Cognition, Quality of Life, Physical Health and Functional

Ability in Alzheimer's Disease: The Effect of Physical Exercise (ADEX Trial): Rationale and Design. Neuroepidemiology 2013; 31:198-207.

[KH15] Hoffmann K, Sobol N, Beyer N, Frederiksen KS, Vogel A, Vestergaard K, Brændgaard H, Gottrup H, Lolk A, Jakobsen S, Laugesen L, Gergelyffly R, Hoegh P, Bjerregaard E, Siersma V, Andersen B, Johannsen P, Cotman C, Waldemar G, Hasselbalch S: Moderate to High Intensity Physical Exercise in Patients with Alzheimer's Disease. (submitted).

[RJ07] Johnson RA, and Wichern DW: Applied multivariate statistical analysis 6th Edition. Englewood Cliffs, NJ: Prentice hall, 2007.

[EL15] Lindemer ER, Salat DH, Smith EE, Nguyen K, Fischl B, Greve DN: White matter signal abnormality quality differentiates mild cognitive impairment that converts to Alzheimer's disease from nonconverters. Neurobiology of Aging 2015;xxx;1-15.

[KL09] Leemput KV, Bakkour A, Benner T, Wiggins G, Wald LL, Augustinack J, Dickerson BC, Golland P, Fischl B:Automated Segmentation of Hippocampal Subfields From Ultra-High Resolution In Vivo MRI. Hippocampus 2009;19;549-557.

[GM84] McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM: Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology 1984;34:939–944.

[HM47] Mann HB and Whitney DR: 'On a test of whether one of two random variables is stochastically larger than the other', Annals of Mathematical Statistics 1947;18;50-60.

[IM12] Meier IB, Manly JJ, Provenzano FA, Louie KS, Wasserman BT, Griffith EY, Hector JT, Allocco E, Brickman AM: White matter predictors of cognitive functioning in older adults. J Int Neuropsychol Soc. 2012;May;18(3):414-427.

[JM99] Mu JS, Li WP, Yao ZB, Zhou XF: Deprivation of endogenous brain-derived neurotrophic factor results in impairment of spatial learning and memory in adult rats. Brain Res 1999;835:259–265.

[KM95] Korte M, Carroll P, Wolf E, Brem G, Thoenen H, Bonhoeffer T: Hippocampal long-term potentiation is impaired in mice lacking brain-derived neurotrophic factor. Proc Natl Acad Sci USA 1995;92: 8856–8860.

[MM13] Marlatt MW, Potter MC, Bayer TA, van Praag H, Lucassen PJ: Prolonged Running, not fluoxetine Treatment, Increases Neurogenesis, but does not alter neuropathology, in the 3xTg Mouse Model of Alzheimer's Disease. Curr Top Behav Neurosci. 2013;15:313-40.

[KN09] Nichol K, Deeny SP, Seif J, Camaclang K, Cotman CW: Exercise improves cognition and hippocampal plasticity in APOE ε4 mice. Alzheimers Dement 2009; 5:287–294.

[HO04] Oja, Hannu; Randles, Ronald H. Multivariate Nonparametric Tests. Statist. Sci. 19 2004;4;598-605.

[HP05] van Praag H, Shubert T, Zhao C, Gage FH: Exercise enhances learning and hippocampal neurogenesis in aged mice. J Neurosci 2005;25:8680–8685.

[MR12] Reuter M, Schmansky NJ, Rosas, HD, Fischl B: Within-Subject Template Estimation for Unbiased Longitudinal Image Analysis. Neuroimage 2012;61(4):1402-1418.

[WR84] Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's Disease. Am J Psychiatry 1984;141(11); 1356-1364.

[RR11] Ruscheweyh R, Willemer C, Kruger K, Duning T, Warnecke T, Sommer J, Volker K, Ho HV, Mooren FG, Knecht S, Floel F: Physical activity and memory functions: An interventional study. Neurobiol Aging. 2011;32(7):1304-1319.

[AS82] Smith A. Symbol Digits Modailities Test (SDMT) Manual (revised). Los Angeles: Western Psychological Services; 1982.

[BS14] Segura B, Baggio HC, Marti MJ, Valldeoriola F, Compta Y, Garcia-Diaz AI, Vendrell P, Bargallo N, Tolosa E, Junque C.: Cortical thinning associated with mild cognitive impairment in Parkinson's disease. Mov Disord. 2014 Oct;29(12):1495-503

[YS11] Shim YS, Youn YC, Na DL, Kim SY, Cheong HK, Moon SY, Park KW, Ku BD, Lee JY, Jeong JH, Kang H, Kim EJ, Lee JS, Go SM, Kim SH, Cha KR, Seo SW: Effects of medial temporal atrophy and white matter hyperintensities on the cognitive functions in patients with Alzheimer's disease. Eur Neurol. 2011;66(2):75-82.

[AV12] Vreugdenhil A, Cannell J, Davies A, Razay G: A community-based exercise programme to improve functional ability in people with Alzheimer's disease: a randomized controlled trial. Scand J Caring Sci. 2012;26(1):12-19.

[JW07] Whitwell J, Przybelski S, Weigand SD, Knopman DS, Boeve BF, Petersen RC, Jack CR: 3D Maps from Multiple MRI Illustrate Changing Atrophy Patterns as Subjects Progress from MCI to AD. BRAIN. 2007;130(7):1777-1786.

Supplementary

Table 4: Gyri and sulci according to known progression of AD partitioned into categories of 'early', 'middle', 'late' and 'very late'.

Gyri	
Early	temporal (inferior, medial, superior lateral), temporal (transversal, plan-polar, plan-tempo), precuneus, cingulate (transversal+ventral)
Middle	parietal (inferior-angular+supramar, superior), occipital-temporal medial (parahippocampal, fusiform)
Late	frontal (inferior-opercular/orbital/triangular, medial, superior)
Very Late	precentral, postcentral
Sulci	
Early	temporal (inferior, superior, transverse), cingulate-marginalis
Middle	parieto-occipital, occipital-temporal lateral, occipital (anterior, middle-lunatus, superior-tranversal)
Late	frontal (inferior, middle, superior)
Very Late	precentral (inferior, superior), postcentral

Cortical regions (Freesurfer, Destrieux atlas nomenclature)

Table 5: Gyri and sulci of the frontal and cingulate cortex, as well as the precentral and postcentral cortex.

Effect	frontal (inferior, middle, superior) gyri and sulci, cingulate (mid-posterior, mid-anterior, anterior) gyri and sulci
Control	precentral gyri, postcentral gyri, postcentral sulci, precentral (inferior, superior) sulci

Results, additional tables

Table 6: Correlations between baseline normalized WMH and change in volume in the hippocampal subfields, parahippocampus, caudate and putamen for all study participants.

		Left side		Right side	
		r-value	p-value	r-value	p-value
Test 1	***Hippocampus***				
	Sum of all subfields	-0.263	0.106	0.147	0.373
	Presubiculum	**-0.345**	**0.031**	0.238	0.145
	Subiculum	-0.295	0.068	0.263	0.106
	Fimbria	0.281	0.083	0.009	0.956
	Hippocampal fissure	-0.225	0.169	-0.313	0.053
	CA1	-0.098	0.552	0.058	0.726
	CA2_3	0.042	0.802	0.085	0.607
	CA4_DG	**-0.406**	**0.010**	0.132	0.423
Test 2					
	Parahippocampus	-0.058	0.721	0.135	0.407
Test 3					
	Caudate	0.301	0.059	0.204	0.208
	Putamen	-0.140	0.390	0.012	0.941

Table 7: correlations between changes in caudate, putamen volume/frontal, cingulate cortical thickness and changes in SDMT/VFT/Stroop outcome measures for all participants.

		Effect (p-value)	Control (p-value)
Frontal, cingulate cortex			
	SDMT	**0.0246**	0.6749
	Stroop	0.8873	0.5966
	VTF	**0.0259**	0.4788
Caudate, Putamen			
	SDMT	0.7042	0.1978
	Stroop	0.8424	0.7530
	VTF	0.0801	0.3846

Table 8: Correlations between relative changes in frontal/cingulate cortical thickness and VFT/SDMT outcome measures for all participants.

	VFT (r-value)	VFT (p-value)	SDMT (r-value)	SDMT (p-value)
Left:				
Frontal inferior sulci	0.037	0.825	0.103	0.549
Frontal middle sulci	0.182	0.275	0.187	0.275
Frontal superior sulci	0.249	0.132	0.099	0.566
Frontal inferior-opercular gyri	0.297	0.071	-0.069	0.690
Frontal inferior-orbital gyri	0.164	0.326	-0.057	0.743
Frontal inferior-triangular gyri	0.319	0.051	-0.032	0.854
Cingulate mid-posterior gyri and sulci	**0.384**	**0.017**	0.118	0.494
Right:				
Frontal inferior sulci	0.174	0.297	0.167	0.324
Frontal middle sulci	0.276	0.093	-0.179	0.296
Frontal superior sulci	0.227	0.170	0.208	0.224
Frontal inferior-opercular gyri	0.269	0.103	0.279	0.104
Frontal inferior-orbital gyri	0.114	0.494	**0.464**	**0.004**
Frontal inferior-triangular gyri	0.277	0.092	**0.386**	**0.020**
Cingulate mid-posterior gyri and sulci	**0.328**	**0.044**	-0.126	0.463

Significant p-values have been highlighted in orange, bold and the corresponding correlations in blue, bold. P-values for the individual gyri and sulci regions have not been corrected for multiple comparisons.

Appendix E

Software-Package for Automated Bias-Field Correction and Quantifcation of Abdominal Fat using MRI with Preliminary Evaluation in Overweight Subjects

Technical Report - DTU Compute

Software-Package for Automated Bias-Field Correction and Quantification of Abdominal Fat using MRI with Preliminary Evaluation in Overweight Subjects

Anders Nymark Christensen, Christian Thode Larsen,
Martin Bæk Petersen, Knut Conradsen, Vedrana Dahl Andersen

December 1, 2015

1 Abstract

The abdominal distribution of fat is investigated in many clinical studies. A major limitation is the time-consuming quantification of the fat, if done manually. We present an automatic and robust package for MRI, that performs both bias-correction and segmentation of both the sub-cutaneous fat and visceral fat. The method uses various image-derived energies, and by tuning them it can be applied to other body parts, such as the thighs. Preliminary assessment on overweight subjects showed few errors when evaluated by a trained medical doctor.

2 introduction

The abdominal distribution of fat is of great clinical interest. Multiple studies [Miyazaki et al., 2002, Fujioka et al., 1987, Rosito et al., 2008] has shown a correlation between abdominal fat, insulin resistance and other metabolic risk factors. The longitudinal changes has also been studied, in e.g. children [Tinggaard et al., 2015] and pre-/post menopausal women [Kanaley et al., 2001, Franklin et al., 2009].

To analyse the distribution of abdominal fat three techniques have seen wide spread use. Dual X-ray absorptiometry (DEXA), Computed Tomography (CT), and Magnetic resonance Imaging (MRI). DEXA is fast and cheap, and only has a very low dose of ionising radiation. However, it does not allow for differentiation between sub-cutaneous fat and visceral fat. CT is fast and gives quantitative measures making the post-processing easier. The drawback is, that it incurs a significant radiation dose. MRI is slow and more difficult to analyse, as it is not quantitative and suffers from a bias-field. As it is free of ionising radiation and has superior contrast it is often the chosen technique. It is also the technique evaluated in this report.

Several algorithms for abdominal fat segmentation has been proposed. SliceOmatic (Tomovision, Inc., Magog, Canada) is a commercial program that allows for segmentation in a supervised fashion. [Armao et al., 2006] used region growing with mixed results. [Leinhard et al., 2008] used an automatic morphon based approach [Knutsson and Andersson, 2005] on a Dixon sequence [Dixon, 1984, Ma, 2008]. While the results are very convincing the article is very sparse with regards to implementation details. Further, it is unlikely that an atlas based on a single subject, as they use, will perform reliable across the wide variety seen in the population. [Zhou et al., 2011] proposes supervised segmentation based on chain-coding [Baruch and Loew, 1988]. Unfortunately the details are omitted for a later publication. [Thörmer et al., 2013] presents a software package available on request for supervised segmentation. The method is based on localised region-based contous [Lankton and Tannenbaum, 2008]. [Liou et al., 2006] proposed an automatic method using binary operations with added heuristics for segmentation. It seems a promising methods, but does not differentiate between visceral fat and internal non-visceral compartments. [Kullberg et al., 2007] proposed a automatic heuristical method that requires the scan to be centred on the L4-L5 vertebra. [Mosbech et al., 2011] used active contours for automatic segmentation. Further, dynamic programming was used to find the fascia of Scarpa, dividing the subcutaneous fat in a deep and superficial part.

We present an integrated package that contains both fully automated bias-

correction and segmentation. The bias field correction method was previously presented in [Larsen et al., 2014], and has been found to perform well across a range of sequences. For the segmentation we implemented the graph-cut based method proposed by [Li et al., 2006], which can be used to perform a global segmentation, thus taking advantage of information in adjacent slices. The method is robust and by weighting different energies, it can be made to work across different sequences and even different body parts. For the exclusion of non-visceral fat some information of the anatomical location of the slice is needed. Except for that it is fully automatic.

3 Materials and Methods

All code was written in Matlab (The MathWorks Inc., Natick, USA).

3.1 Bias-correction

The bias field correction method employs a generative model, where the bias field is assumed to be multiplicative, smooth and slowly varying over the image. The generative model is composed of a mixture of Gaussians that models the underlying uncorrupted signal intensities and a linear combination of cubic b-spline basis functions with regularization on the bending energy for the bias field effect.

Parameter estimation of the generative model is carried out in the log-transformed data domain using a generalized expectation maximization algorithm [Van Leemput et al., 1999]. The following hyper-parameters were used: 8 Gaussians, smoothing distance 25, smoothing regularisation 0.5, as they were found to perform well across a range of MRI sequences.

The implementation of our generative bias field correction method was previously presented in [Larsen et al., 2014]. Here it was shown how performance of the method is comparable to the popular N3 bias field correction method [Sled et al., 1998], and also how it is important that involved hyper-parameters are tuned properly, depending on scanner and field strength.

3.2 Segmentation

The segmentation of the images is based on the method by [Li et al., 2006], were the problem is formulated as a graph to be cut. The scan is unrolled for

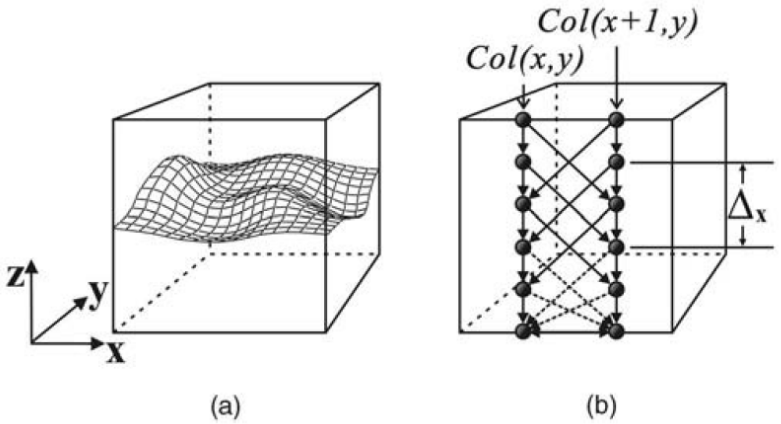


Figure 1: (a) Surface orientation. (b) Two adjacent columns with edges.

the segmentation. If we have an orientation as in fig. 1 (a), the surface will only pass through each (x,y)-column once. To ensure connectivity we have intra-column edges in each column and to incorporate smoothness constraints Δ_x, Δ_y we have inter-column edges, see fig. 1 (b). This can be generalised to 4D if we have multiple surfaces and want to impose constraints Δ_l, Δ_u on the distance between them. The node weights (energy used) are set depending on the problem. If it is layers in a 3D volume, it could e.g. be the inverse of the gradient, as we cut in the minimum value. To cut the graph we use [Boykov and Kolmogorov, 2004].

3.2.1 Subcutaneous fat

The inner and outer surface of the subcutaneous fat are found together, with $\Delta_l = 5$, $\Delta_u = 150$. The volume needs to be unrolled. This is done by finding the centroid and then sample radially. As the inner surface of the subcutaneous fat changes distance to the centroid significantly in the lateral areas, we have a denser sampling in those. See fig. 2. The resulting unrolled image, see fig. 3, can then be used to estimate energies for the segmentation.

Different energies are used to find the inner and outer surface. A gradient based G, a gradient based of the median filtered scan G_m, a gradient based

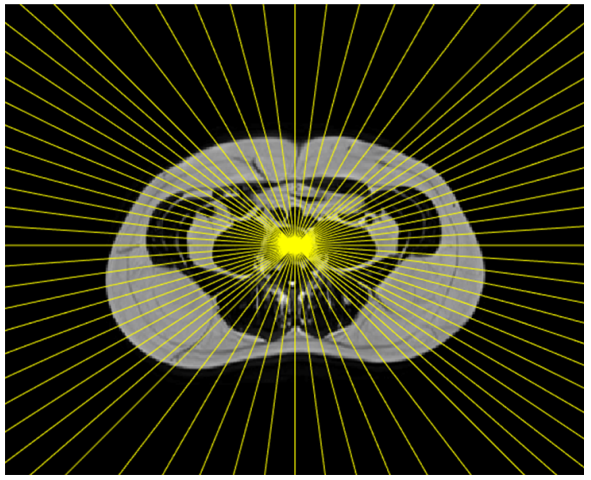


Figure 2: Radial sampling

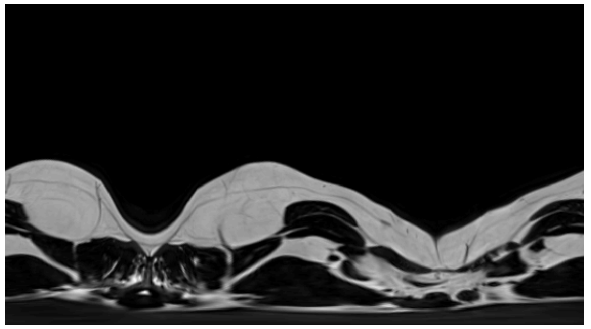


Figure 3: Unrolled image

of the median filtered scan with additional Gaussian smoothing G_s and a surface cost S.

With the slices along the x-direction, the gradient based energy is defined as

$$G = \text{sign}(G_z) \cdot (|G_z| + |G_x|) \tag{1}$$

Where G_z and G_x are the gradients in the z- and x-direction.

The gradient of the median filtered scan G_m is a more robust measurement. It is obtained by median filtering the scanning with a kernel of size round(10/pixel size in mm) and then taking the gradient measure.

The gradient of the median filtered scan with additional Gaussian smoothing G_s used the same kernel as above for the median filter and for the Gaussian a kernel of size 1.5 times the one used in median smoothing with a standard deviation of 3/pixel size in mm.

The surface cost S is a product of two energies. The first S_1 is given by setting all values in G less than zero to zero and then taking the cumulative sum in the z-direction. The second S_2 is given by taking the cumulative sum in z-direction of the unrolled image. The two energies are normalised and multiplied. This energy is then filtered with a median filter with the kernel [round(70/pixel size in mm) round(4/pixel size in mm)] (x,z).

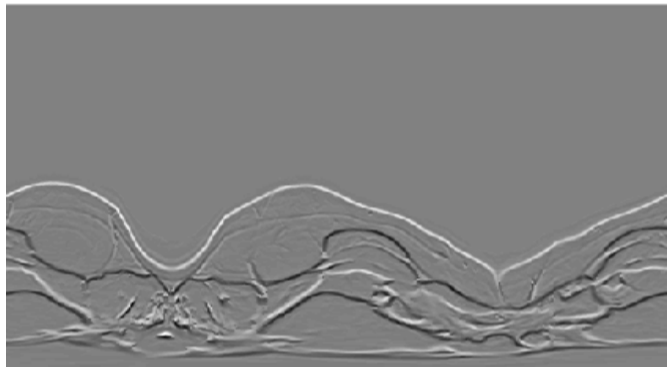
To find the outer surface normally only G_s is used, but it can be weighted with the other energies if necessary. The values Δ_x, Δ_y are set to 10 and 10. For the inner surface a weighting of G, G_m, and S is used, specifically:

$$E_i = 0.5 \cdot G + 0.1 \cdot G_m + 0.4 \cdot S \tag{2}$$

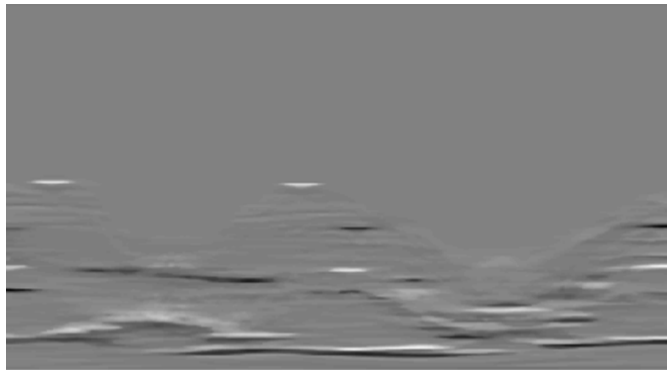
The found inner surface is median filtered with a kernel of 5, except in the anterior region were a kernel of 35 was used. This was due to problems with visceral fat being segmented as subcutaneous. An example of the obtained segmentation is shown in fig. 6. The values Δ_x, Δ_y are set to 3 and 7.

3.2.2 Visceral fat

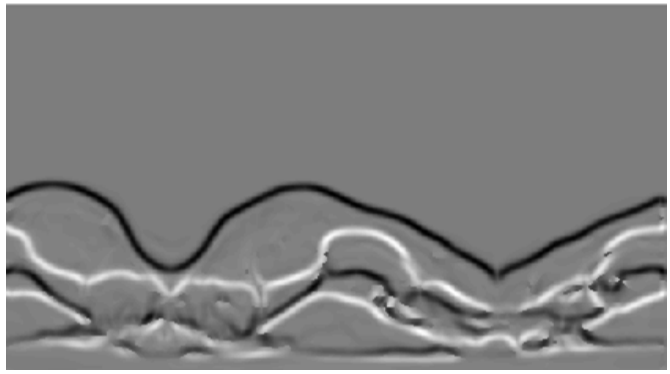
Using the inner surface of the subcutaneous fat as a mask, the visceral fat was segmented. We use a threshold value based on k-means clustering with 5 classes. Ordering the clusters after mean value in ascending order, we used the value of the 4th cluster as threshold. Empirically we discovered that the threshold obtained by using Otsu threshold [Otsu, 1979] were too low, see fig. 7.



(a) G

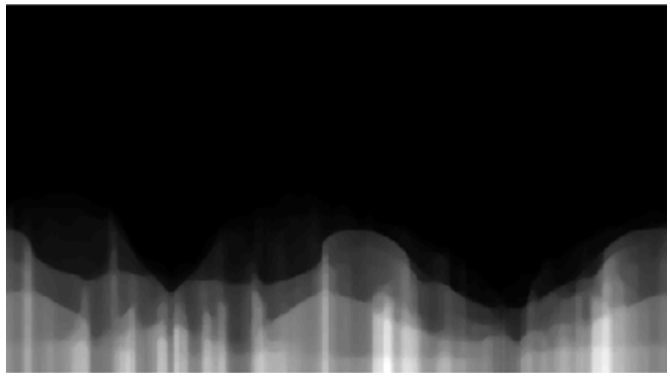


(b) G_m

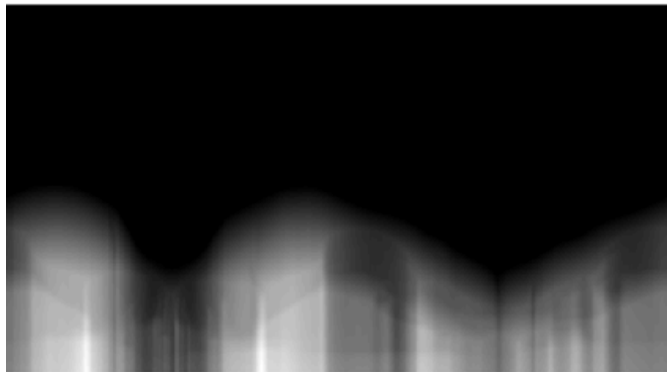


(c) G_s

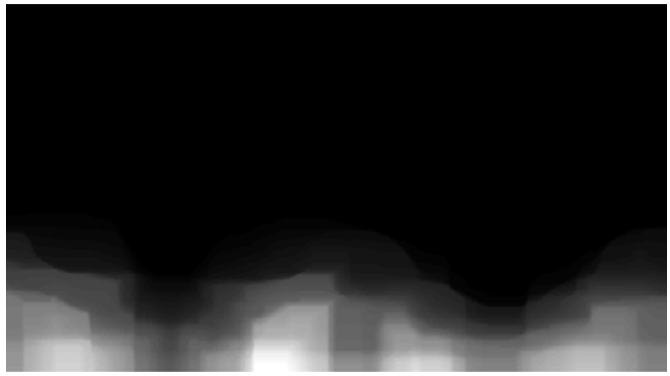
Figure 4: The three gradient based energies



(a) S_1

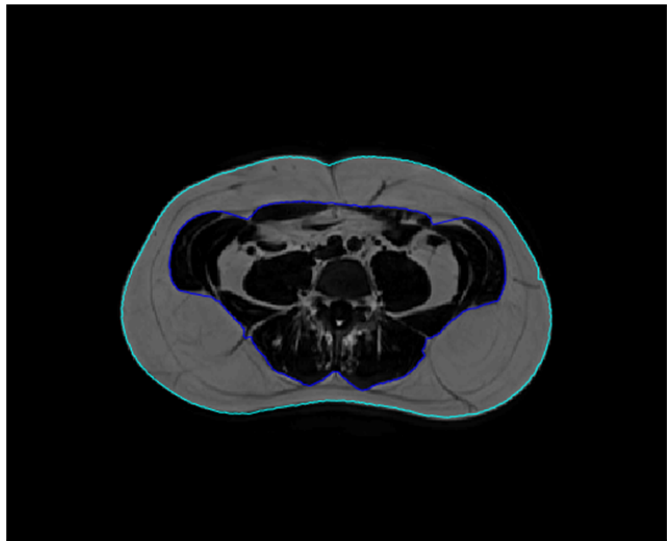


(b) S_2

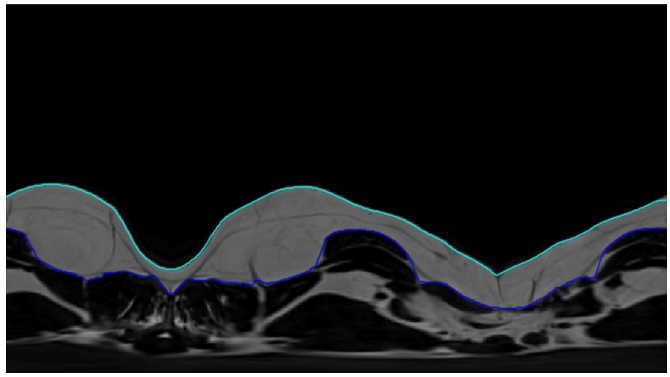


(c) S

Figure 5: The three surface cost based energies



(a) Original



(b) Unrolled

Figure 6: Segmentation of the subcutaneous fat

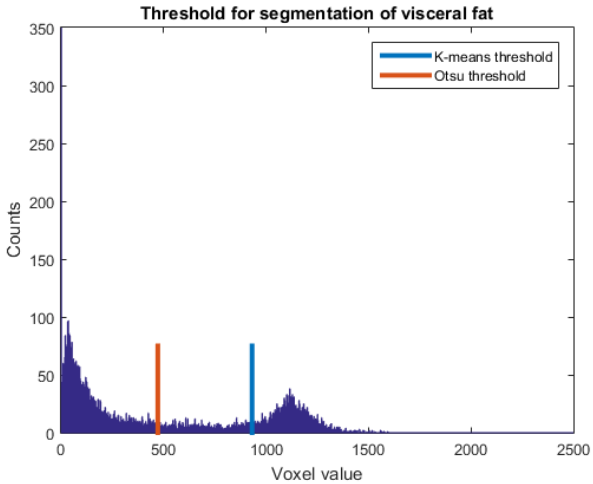


Figure 7: Histogram showing the difference between Otsu and k-means threshold

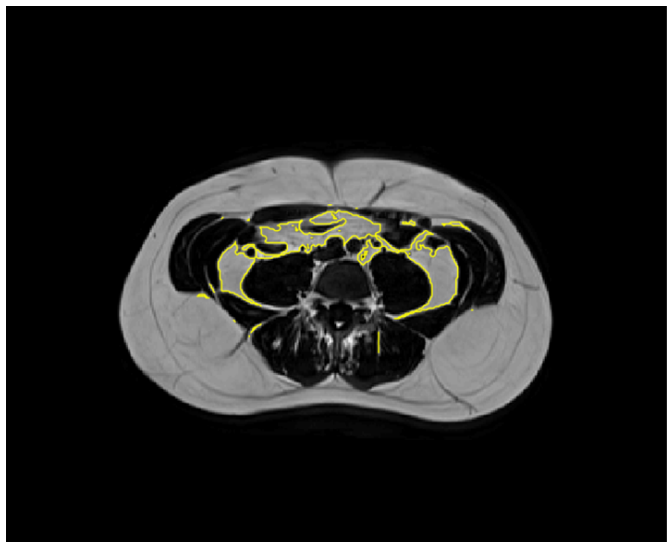


Figure 8: The visceral fat, with the spinal column area extracted

3.2.3 Spine extraction

The spine extraction is optional. Without it, the fat in the non-visceral compartment around the spine will be classified and counted as visceral fat. It needs some user input, as it needs to be centred approximately on the spine. In the lover parts of the spinal column it is nearly centred in the slice. When moving caudally the spine shifts more posterior. The user can input either a vector describing the displacement or a single number used for all slices.

Centred on the spine the scan is unfolded radially as before. The energies, however, are different. For all of them, except the surface cost, the anterior part (1/4 radially) of the scan was set to zero.

The first is simply the raw image im. The second is the gradient of the raw image G^{spine}. The third is the gradient of the scan filtered with a median filter G_m^{spine} using a kernel of size round(10/pixel size in mm). The surface costs are a bit more complicated. The cumulative sum is taking in the z-direction. For the first S_1^{spine} starting from the bottom, for the second S_2^{spine} starting from the top. For S_1^{spine} the lower 1/30 of the z-axis are set to infinity. For S_2^{spine} the upper 3/4 are set to infinity. This constrains the possible solution.

The weighting of the different energies is as follows:

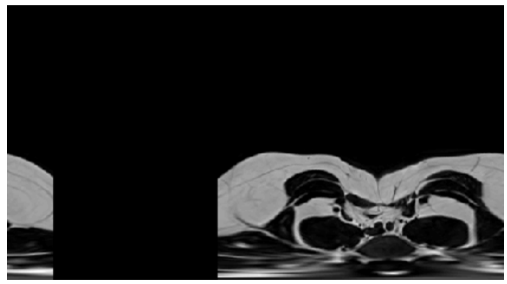
$$E_{spine} = 0.1 \cdot im + 0.2 \cdot G^{spine} + 0.25 \cdot G_m^{spine} + 0.4 \cdot S_1^{spine} + 0.1 \cdot S_2^{spine} \quad (3)$$

3.3 Test subjects

10 healthy subjects included in the project 'Active Commuting To Improve Wellbeing and health in Everyday life' (ACTIWE) was randomly chosen. The subjects were 20-45 years of age and overweight with a BMI of 25-35. Imaging was performed on a Biograph mMR (Siemens, Erlangen, Germany) using a T1 weighted sequence with water suppression. 5 slices centred on the L5 vertebra were chosen for analysis. The slices were segmented, and an experienced medical doctor (MD) manually annotated any errors.

3.4 Earlier studies

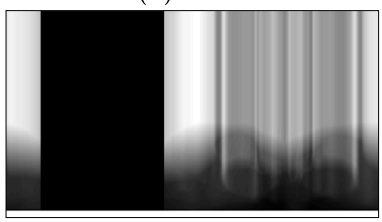
During the development of this software, 2 other projects have been analysed.



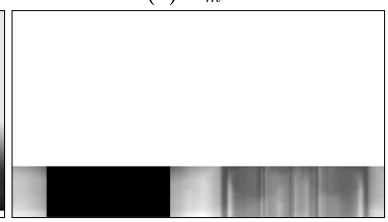
(a) im

(b) G^{spine}

(c) G_m^{spine}

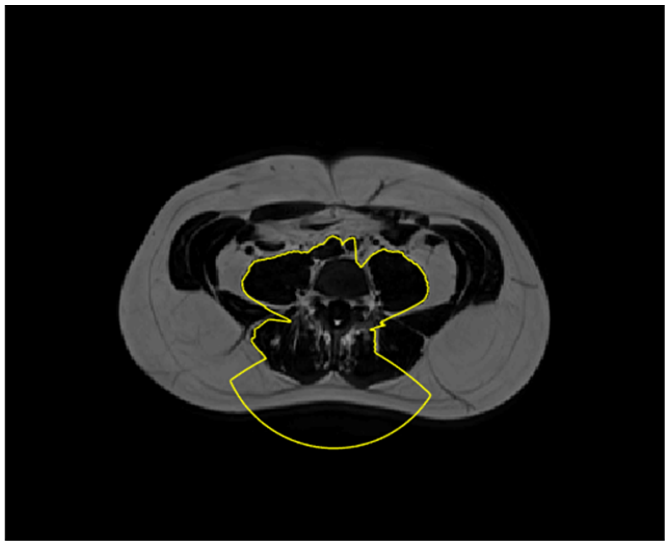


(d) S_1^{spine}

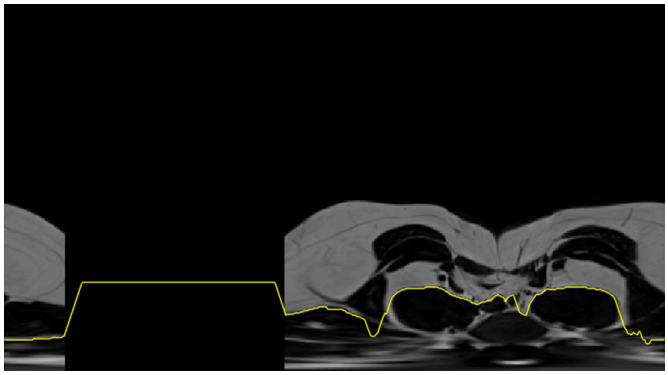


(e) S_w^{spine}

Figure 9: The energies used for the spine extraction



(a) Original



(b) Unrolled

Figure 10: Extraction of the spine

3.4.1 MRI scans of children using Dixon

197 children between 10 and 15 years of age were examined in a longitudinal study at Department of Growth and Reproduction, Copenhagen University Hospital Rigshospitalet. The imaging sequence used was Dixon.

The bias-correction was performed using an older method [Mosbech et al., 2011]. The segmentation was performed using an early iteration of the method described here. A single study was re-examined, to evaluate the robustness of our software package on different patient groups and imaging sequences.

3.4.2 Abdominal and thigh scans of adult men using T1

In the study 'Testosteron treatment of men with type 2 diabetes' conducted at Department of Endocrinology at Odense University Hospital segmentation was performed on both abdominal slices and thigh slices. 27 men were scanned twice and imaging were recorded using a T1 sequence without water suppression.

The analysis was also performed using an earlier iteration of the method described here. A single study was reexamined to evaluate the robustness of our software package on different patient groups, imaging sequences and body parts.

4 Results

The segmentation results evaluated by a MD is presented in table 1. The error was evaluated with Matthews Correlation Coefficient [Matthews, 1975] and The Dice coefficient. The overall misclassification was 0.233 % and 0.101 % for visceral and subcutaneous fat respectively.

A number of repeated errors were identified. For the subcutaneous fat, the outer surface was placed inside the fat in a right lateral position. The error was minor but systematic. Anterior laterally there was often a minor error at the intersection of the subcutaneous fat, the abdominal compartment and the lateral muscle compartments. A larger error but not as common, was a chunk of visceral fat being classified as subcutaneous in the anterior part.

For the visceral fat, the spine extraction has to some degree errors. These errors go both way, both too much and too little extracted. Further, the visceral fat seems to be underestimated. The false positives are predominantly from the lateral muscle compartments.

Table 1: Segmentation results from the ACTIWE study. TP = True Positive, TN = true Negative, FP = False Positive, FN = False Negative, MCC = Matthews Correlation Coefficient, DICE = Dice coefficient

Visceral fat						Subcutaneous fat					
TP	TN	FP	FN	MCC	DICE	TP	TN	FP	FN	MCC	DICE
34154	564085	497	304	0.988	0.988	48230	550508	197	105	0.997	0.997
57548	540147	890	455	0.987	0.988	55125	542257	1203	455	0.984	0.985
45159	552659	770	452	0.986	0.987	90325	508564	0	151	0.999	0.999
10253	587165	754	868	0.925	0.927	71477	527469	68	26	0.999	0.999
20920	577270	595	255	0.979	0.980	98872	499985	136	47	0.999	0.999
54534	542924	1255	327	0.984	0.986	66935	532050	0	55	1	1
34358	563359	1202	121	0.980	0.981	48936	548819	437	848	0.986	0.987
45115	553113	658	154	0.990	0.991	87724	510035	1044	237	0.992	0.993
17371	579254	1947	468	0.934	0.935	57395	540716	312	617	0.991	0.992
17809	579269	1313	649	0.946	0.948	69356	529544	0	140	0.999	0.999

The analysis of a scan of a child using Dixon, is shown in fig. 11. The bias-correction has worked fine, as well as the segmentation of the subcutaneous fat. The visceral fat is, however underestimated.

The analysis of a man in the testosteron study is shown in fig. 12. No problems were encountered with the bias-correction. For the thigh it was necessary to change the outer surface energy to

$$E_o = 0.75 \cdot G_s + 0.25 \cdot S_2 \tag{4}$$

as the femur had a very strong gradient. Some errors are apparent in the outer surface, overestimation in the top, and a minor underestimation to the right. The inner surface lacks support in the lower right part, but except for that performs acceptable.

For the abdominal slices the Δ_y for the inner surface was changed to 70, as the slices were very far apart and changed considerably. The spine extraction has failed, and is apparently sensitive to changes in sequence. Except for that the segmentation is satisfactory.

5 Discussion

We have presented a fully automatic bias-correction and segmentation method that works robustly across different sequences.

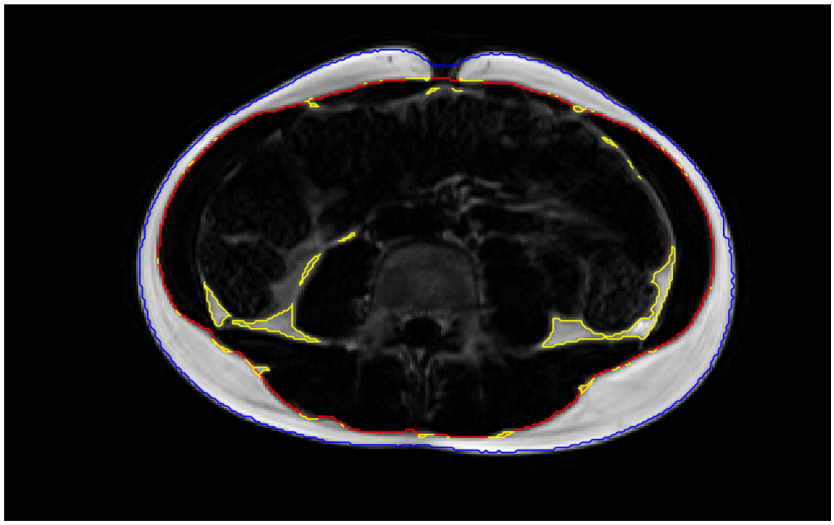
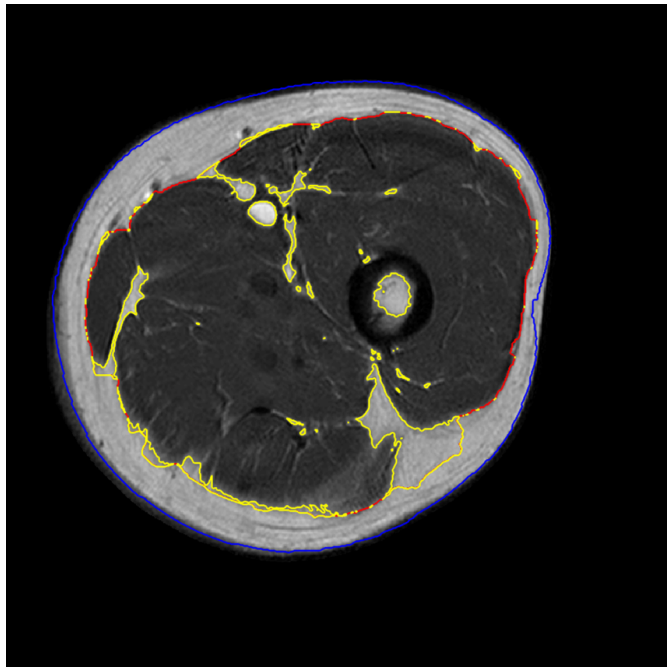


Figure 11: Scan of a child using Dixon

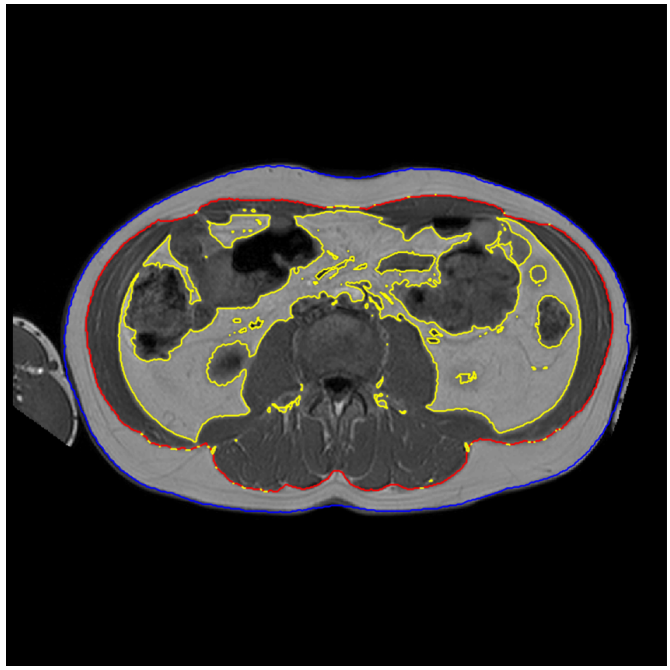
As seen from the results the method works very well in general. Our validation method where a MD marks errors is likely to be biased to some degree, as the answer is already given by the scan. However, using blind segmentations is not without problems, as this suffers from both intra and inter observer variance. Even with these drawbacks blinded manual segmentation is often used as the 'ground truth' in comparisons.

This said the bias-correction and segmentation works well, but some problems persist. Sometimes the visceral fat and subcutaneous fat present with no visible demarcation. To fix that a heavier smoothing of the inner surface is used in the anterior part, but even that is not sufficient to adequately correct the problem. The threshold for the visceral fat may need to be tuned slightly downwards and certainly needs to be sequence dependant. The current threshold works well in T1 and T1 with water suppression, but not in Dixon. The original analysis of the Dixon images used Otsu threshold and this could easily be incorporated. The spine extraction is rather complicated and not sufficiently robust. In fig. 10 it misses the right erector spinae muscle compartment. In the T1 images it fails completely. In many studies the non-visceral fat is, however, counted as visceral and depending on the study the lack of this feature may not be critical.

An issue not mentioned before is the lateral muscle compartments. While



(a) Thigh



(b) Abdomen

Figure 12: Segmentation of a man from the testosteron study using T1

they normally contain very little fat, they should ideally be classified as a non-visceral compartment, and the fat present in them counted separately.

6 Conclusion

We have presented a robust, easy to use routine, that performs all the necessary steps for segmentation of abdominal fat. Some work remains before the software is ready to be made public. See future work below.

7 Future Work

In addition to a rigorous validation across different sequences and population groups, some features should be incorporated.

7.1 Fascia of Scarpa

The fascia of Scarpa partitions the subcutaneous fat. The blood from the inner part drains to the liver, and it is hypothesised that it is thus too simplified to look at the subcutaneous fat as one compartment. The fascia of Scarpa is fickle and not easily identified on many slices. Some work has been done see e.g. [Mosbech et al., 2011]. The segmentation of this fascia should be incorporated in the package although it is only feasible in T1 with water suppression and maybe T1 images. A significant problem is how to handle the segmentation in the slices with no support. Especially in the anterior part, it is often impossible to identify the fascia.

7.2 Non-visceral compartments

The spine extraction needs to be more robust. Moreover, the lateral muscle compartments should also be extracted and counted separately.

7.3 Automatic spine detection

Instead of relying on user input the spine should be identified automatically from the scan. As it has a specific repeating structure it should be possible, but may be quite challenging.

7.4 Modify the graph-cut problem

Especially in the anterior part, visceral fat is often located adjacent to the subcutaneous fat, without a visible demarcation on the scans. By making the constraints on the graph-cut radially dependant, it would be easier to achieve a better segmentation in those cases.

7.5 Processing time

The processing time, can be around half an hour for larger scans with around 40 slices. This should be brought down, e.g. by cropping the data-set and excluding the air voxels around the subject. This would also reduce the memory requirements.

7.6 Bias-correction

Needs to be validated rigorously across the different sequences. Although problems have not been encountered, it is probable that a lower number of Gaussians will perform more robustly.

References

[Armao et al., 2006] Armao, D., Guyon, J. P., Firat, Z., Brown, M. a., and Semelka, R. C. (2006). Accurate quantification of visceral adipose tissue (VAT) using water-saturation MRI and computer segmentation: Preliminary results. *Journal of Magnetic Resonance Imaging*, 23(5):736–741.

[Baruch and Loew, 1988] Baruch, O. and Loew, M. H. (1988). Segmentation of two-dimensional boundaries using the chain code. *Pattern Recognition*, 21(6):581–589.

[Boykov and Kolmogorov, 2004] Boykov, Y. and Kolmogorov, V. (2004). An experimental comparison of min-cut/max- flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(9):1124–1137.

[Dixon, 1984] Dixon, W. T. (1984). Simple proton spectroscopic imaging. *Radiology*, 153(1):189–194.

[Franklin et al., 2009] Franklin, R. M., Ploutz-Snyder, L., and Kanaley, J. a. (2009). Longitudinal changes in abdominal fat distribution with menopause. *Metabolism: Clinical and Experimental*, 58(3):311–315.

[Fujioka et al., 1987] Fujioka, S., Matsuzawa, Y., Tokunaga, K., and Tarui, S. (1987). Contribution of intra-abdominal fat accumulation to the impairment of glucose and lipid metabolism in human obesity. *Metabolism: clinical and experimental*, 36(1):54–59.

[Kanaley et al., 2001] Kanaley, J. a., Sames, C., Swisher, L., Swick, a. G., Ploutz-Snyder, L. L., Steppan, C. M., Sagendorf, K. S., Feiglin, D., Jaynes, E. B., Meyer, R. a., and Weinstock, R. S. (2001). Abdominal fat distribution in pre- and postmenopausal women: The impact of physical activity, age, and menopausal status. *Metabolism: Clinical and Experimental*, 50(8):976–982.

[Knutsson and Andersson, 2005] Knutsson, H. and Andersson, M. (2005). Morphons: Segmentation using elastic canvas and paint on priors. *Proceedings - International Conference on Image Processing, ICIP*, 2:1226–1229.

[Kullberg et al., 2007] Kullberg, J., Ahlström, H., Johansson, L., and Frimmel, H. (2007). Automated and reproducible segmentation of visceral and subcutaneous adipose tissue from abdominal MRI. *International journal of obesity (2005)*, 31(12):1806–1817.

[Lankton and Tannenbaum, 2008] Lankton, S. and Tannenbaum, A. (2008). Localizing region-based active contours. *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, 17(11):2029–39.

[Larsen et al., 2014] Larsen, C. T., Iglesias, J. E., and Van Leemput, K. (2014). N3 Bias Field Correction Explained as a Bayesian Modeling Method. pages 1–12.

[Leinhard et al., 2008] Leinhard, O., Johansson, a., Rydell, J., Smedby, O., Nyström, F., Lundberg, P., and Borga, M. (2008). Quantitative abdominal fat estimation using MRI. *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, pages 1–4.

[Li et al., 2006] Li, K., Wu, X., Chen, D. Z., and Sonka, M. (2006). Optimal surface segmentation in volumetric images - A graph-theoretic approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(1):119–134.

[Liou et al., 2006] Liou, T.-H., Chan, W. P., Pan, L.-C., Lin, P.-W., Chou, P., and Chen, C.-H. (2006). Fully automated large-scale assessment of visceral and subcutaneous abdominal adipose tissue by magnetic resonance imaging. *International journal of obesity (2005)*, 30(5):844–852.

[Ma, 2008] Ma, J. (2008). Dixon techniques for water and fat imaging. *Journal of Magnetic Resonance Imaging*, 28(3):543–558.

[Matthews, 1975] Matthews, B. W. (1975). Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochimica et biophysica acta*, 405(2):442–451.

[Miyazaki et al., 2002] Miyazaki, Y., Glass, L., Triplitt, C., Wajcberg, E., Mandarino, L. J., and DeFronzo, R. a. (2002). Abdominal fat distribution and peripheral and hepatic insulin resistance in type 2 diabetes mellitus. *American journal of physiology. Endocrinology and metabolism*, 283(6):E1135–E1143.

[Mosbech et al., 2011] Mosbech, T. H., Pilgaard, K., Vaag, A., and Larsen, R. (2011). Automatic segmentation of abdominal adipose tissue in MRI. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6688 LNCS:501–511.

[Otsu, 1979] Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66.

[Rosito et al., 2008] Rosito, G. a., Massaro, J. M., Hoffmann, U., Ruberg, F. L., Mahabadi, A. a., Vasan, R. S., O'Donnell, C. J., and Fox, C. S. (2008). Pericardial fat, visceral abdominal fat, cardiovascular disease risk factors, and vascular calcification in a community-based sample the framingham heart study. *Circulation*, 117(5):605–613.

[Sled et al., 1998] Sled, J. G., Zijdenbos, a. P., and Evans, a. C. (1998). A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE transactions on medical imaging*, 17(1):87–97.

[Thörmer et al., 2013] Thörmer, G., Bertram, H. H., Garnov, N., Peter, V., Schütz, T., Shang, E., Blüher, M., Kahn, T., and Busse, H. (2013). Software for automated MRI-based quantification of abdominal fat and preliminary evaluation in morbidly obese patients. *Journal of Magnetic Resonance Imaging*, 37(5):1144–1150.

[Tinggaard et al., 2015] Tinggaard, J., Hagen, C., Mouritsen, A., Mieritz, M., Wohlfahrt-Veje, C. Fallentin, E., Larsen, R., Christensen, A., Jensen, R., Juula, A., and Maina, K. (2015). Abdominal fat Distribution Measured by Magnetic Resonance Imaging in 197 Children Aged 10–15 Years – Correlation to Anthropometry and Dual X-Ray Absorptiometry. In *ESPE Abstracts*, pages 84 P–2–330.

[Van Leemput et al., 1999] Van Leemput, K., Maes, F., Vandermeulen, D., and Suetens, P. (1999). Automated model-based bias field correction of MR images of the brain. *IEEE transactions on medical imaging*, 18(10):885–896.

[Zhou et al., 2011] Zhou, A., Murillo, H., and Peng, Q. (2011). Novel segmentation method for abdominal fat quantification by MRI. *Journal of Magnetic Resonance Imaging*, 34(4):852–860.

Appendix F

Lung disease classification in mice using x-ray phase-contrast and dark-field imaging

Lung disease classification in mice using x-ray phase-contrast and dark-field imaging

A. N. CHRISTENSEN, A. YAROSHENKO, K. HELLBACH, K. WILLER, L. GROMANN, A. Ö. YILDIRIM, K. CONRADSEN, F. PFEIFFER

***Abstract*—Pulmonary diseases are among the leading causes of lowered life-quality and morbidity worldwide. Using conventional x-ray imaging, common pulmonary disorders like emphysema and fibrosis are difficult to diagnose in their early stages.**

In contrast to conventional x-ray imaging, dark-field is more sensitive to morphological changes in lung. However, due to similar characteristic behavior of x-ray dark-field for both emphysema and fibrosis, it has been questioned if it is possible to correctly classify the disorders.

We present a method based on a combination of absorption and dark-field signals and incorporation of regional information through a healthy lung model that is shown to classify pixel-wise between healthy, emphysema and fibrosis with an accuracy of 84.05 % in a murine model. An easy way to visualize the probabilistic results is presented using a colour coded image. Finally, the possibility of including the phase-contrast signal for a better classification is analysed.

***Index Terms*— Phase contrast x-ray, dark-field, classification, pulmonary disease, fibrosis, emphysema**

I. INTRODUCTION

Lung diseases such as emphysema and fibrosis are among the leading causes of lowered life-quality and morbidity worldwide [1], [2].

Despite substantial histological changes in the architecture of the lungs, these diseases are difficult to diagnose in their early stage using conventional x-ray radiography [3]–[6].
It has been shown in recent studies [7]–[9] that a superior sensitivity for the detection of these maladies can be obtained using the dark-field radiography. This imaging modality can be acquired additionally to differential phase-contrast and conventional absorption imaging if a three grating Talbot-Lau interferometer is introduced into the beam [10], [11]. The dark-field is a measure for the non-resolvable x-ray small-angle scattering occurring at the air-tissue interfaces in the lung, whereas the phase-contrast visualizes the x-ray beam refraction on directly resolved structures.

The morphological changes, associated with the aforementioned disorders lead to a loss of air-tissue interfaces and consequent loss of x-ray dark-field signal [12]. Due to similar characteristic behavior for different disorders of the dark-field it has been questioned whether x-ray grating-based imaging is suitable for correct classification of the different pulmonary disorders. A prior study [8] has demonstrated proof-of-principle feasibility of classification of the disorders globally using shape parameters of the lung in addition to the mean signal values. In this study we report a method to obtain a direct pixel-wise classification probability using only the signal values. In particular, we investigate the differences between healthy, emphysematous and fibrotic lungs and analyze how much we can improve the classification by incorporating local information through a lung model with respect to the global signal values. Finally, we assess whether the phase image holds additional information that helps the classification.

II. METHODS AND MATERIALS

A. Imaging

All animal experiments were approved by the relevant authorities. Emphysema and fibrosis were induced in animals using an orotracheal instillation with a solution of pancreatic elastase or bleomycin in a sterile phosphate-buffered saline, respectively. The successful induction of the disorders was confirmed by histological analysis. The in vivo imaging was performed using a prototype small-animal x-ray phase-contrast and dark-field scanner [13], [14]. The phase-contrast and dark-field radiographies were obtained from 5 images acquired at different positions of the source grating, with an exposure time of 3.3 seconds per step. The x-ray source was operated at 35 kVp and 18 W. The image resolution is approximately 60 µm (10% MTF). During image acquisition the animals were breathing freely after being anesthetized using intraperitoneal injection of medetomidine (500 µg/kg), midazolam (5 mg/kg), and fentanyl (50 µg/kg).

B. Dataset

In total, 31 healthy, 13 fibrotic and 14 emphysematous mice were analyzed. The mice were from prior studies, e.g. [9], [15], [16] . An expert radiologist, experienced with phase-contrast and dark-field imaging, annotated the fibrotic and

A. N. Christensen (e-mail: anym@dtu.dk) and K. Conradsen are with the Department of Computer Science and Applied Mathematics at the Technical University of Denmark.

A. Yaroshenko, K. Willer, L. Gromann, and F. Pfeiffer are with the Lehrstuhl für Biomedizinische Physik, Physik-Department & Institut für Medizintechnik, Technische Universität München.

K. Hellbach is with the Department of Clinical Radiology at the Ludwig-Maximilians-University Hospital Munich.

O. A. Yildirim is with the Institute of Lung Biology and Disease, Helmholtz Zentrum, Munich.

emphysematous mice by marking regions that were very likely to be diseased. The images were partitioned into a training set with 6 healthy, 6 fibrotic and 6 emphysematous mice, and a test set of 6 healthy, 7 fibrotic and 8 emphysematous mice The remaining 19 healthy mice were used to build a model of the healthy mouse lung.

C. Image Processing

1) Active Appearance Model

To segment the lung, we used an active appearance model (AAM) [8], that automatically delineates the lung using the dark-field image, as shown in Figure 1. It then deforms the lung to a standard shape. In this way direct comparisons can be made across different mice.

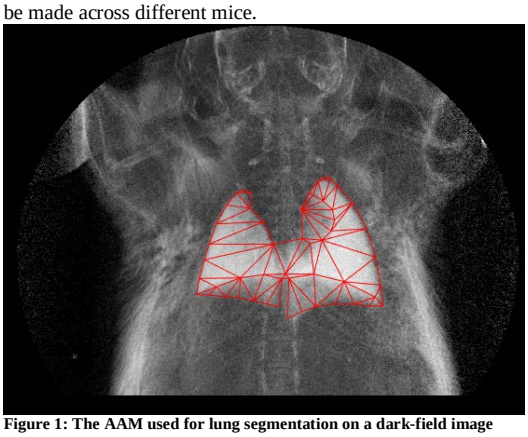


Figure 1: The AAM used for lung segmentation on a dark-field image

2) Rib Extraction

The ribs were masked and excluded from the analysis, since they are visible on the absorption but not on the dark-field image. To find them we used a histogram equalization of the absorption image. The equalized image was filtered by a Laplacian of Gaussian (kernel size: 25x25 pixels, standard deviation of 8 pixels) and a threshold (values above 0.0004) applied. Finally, areas smaller than 500 pixels were removed, see Figure 2.

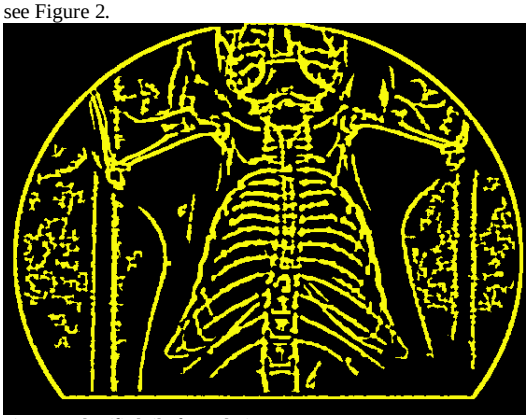


Figure 2: Identified ribs for exclusion

3) Spine extraction

The spine was also masked and excluded from the analysis, see Figure 3. It was localized by median filtering the absorption image, and downsampling it to one fourth of the original size. The gradient was taken in the horizontal direction and added to the image. Finally a graph-cut algorithm [17] was used to find the two minimum value paths from top to bottom of the image, with the constraint that the two paths must have a distance of 20 to 55 pixels between them.

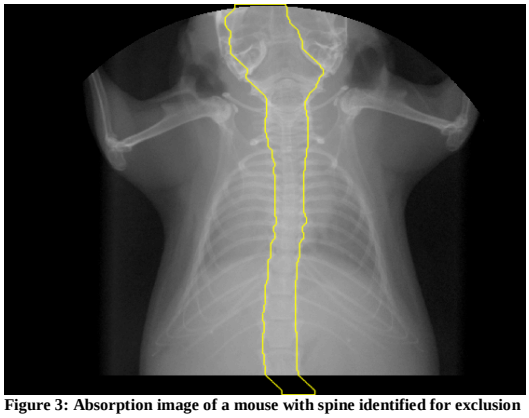


Figure 3: Absorption image of a mouse with spine identified for exclusion

D. Features used

The absorption, dark-field and differential phase images were used directly.

Further, we investigated three possible ways to extract additional information from the differential phase image. These were based on different texture filters: a range filter, a standard deviation filter, and an entropy filter to the image. For each pixel in the image we apply a mask of a certain shape and size. For the range filter, the difference between the maximum and minimum pixel value within the mask is found, and that value is returned to the pixel the mask is centered on. The standard deviation filter returns the standard deviation, while the entropy filter uses the entropy defined by (1).

$$entropy = -\sum_{mask} p \cdot log2(p) \quad (1)$$

where – for each pixel in the mask - p is the probability of the pixel value of the pixel. The probability is estimated as the relative frequency of the pixel value in the entire image.
We used disk shaped masks for all three filters with radius ranging from 3 to 25 pixels. See Figure 4 for an example.

E. Healthy lung model

Using the 19 healthy mice, a model of the healthy lung was generated. The ribs and spine were excluded and the AAM was used for segmentation and deformation to a standard shape. Then the mean and standard deviation in each pixel were calculated across the 19 mice. This was done for the absorption, dark-field, differential phase and the 3 times 23 different images resulting from applying the above filters to the differential phase-contrast image. The healthy lung model

was used as a way to incorporate regional information and thus the varying signal intensity across the lung, see Figure 5.

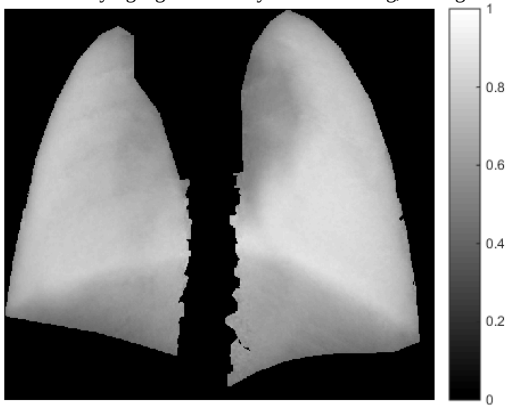


Figure 5: Average dark-field image of the masked lung of 19 healthy mice

F. Histograms

Using the mice from the test set, normalized histograms were calculated for each of the absorption, dark-field, phase and the 3 times 23 different filtered phase images. This gives us 72 different estimated marginal probabilities of a given value belonging to either healthy, emphysematous or fibrotic tissue.

Further we calculated the histograms for the same values but after we have subtracted for each pixel the corresponding healthy lung average and then divided with the pixel standard deviation from the normal lung. An example of the histograms for absorption and dark-field after subtraction and division of the healthy lung can be seen in Figure 6.

G. Classification and model selection

By evaluating the histogram value in each pixel, we can classify by selecting the disease with the highest probability (value in the histogram). When using more than one measure, e.g. both absorption and dark-field, the obtained probabilities from each histogram are simply multiplied. The final classification is then the label: healthy, emphysema or fibrotic, that has the highest probability in a given pixel. Calculated maps for the whole lung can be visualized using a RGB color space, to give a direct impression of how certain a classification is, see Figure 7, 8 or 9. The brighter the color the more certain the classification is and an uncertain classification between e.g. fibrosis and emphysema is highlighted by purple color.

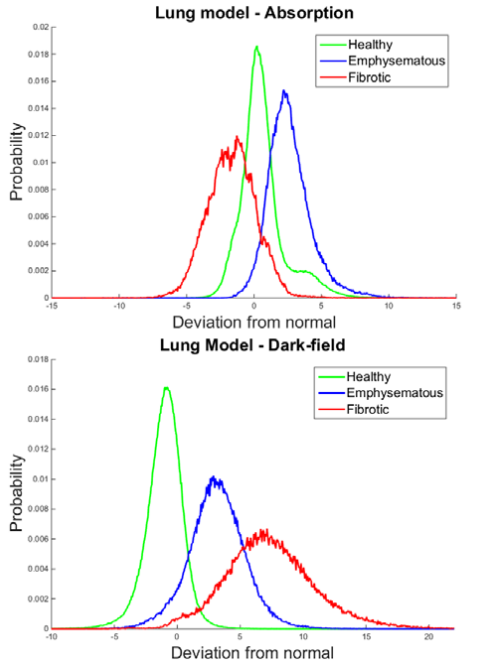


Figure 6: Absorption and dark-field signals histograms obtained from the lung model.

The absorption image, dark-field, differential phase image, and the 23 different filtered images for both range, standard deviation and entropy were tested to see which yielded the best classification.

The model was based on forward selection. All features were tested and the best included. Keeping that chosen feature all the rest was tested and the best included. This procedure was repeated. The measure for success was accuracy (Acc) defined as: how large a fraction of the lungs in the test set that was classified correctly. The accuracy was determined for each mouse in the test set, and the mean of all accuracies was used.

To penalize complex models, the Bayesian information criterion (BIC) [18] and Akaikes information Criterion, corrected for small sample size (AIC) [19] was used.
They are defined respectively:

$$BIC = n \cdot \ln(1 - Acc) + k \cdot \ln(n) \quad (2)$$

$$AIC = 2k + n \cdot \ln(1 - Acc) + \frac{2k(k+1)}{n-k-1} \quad (3)$$

where n is the number of samples (mice), and k the number of parameters in the model (features included). The model yielding the smallest value is then preferred.

III. Results

A. Training set

We tested four separate sets:

- Raw image values: using only the raw values with no lung model.
- Lung model: where the pixel values have been corrected by subtraction of the mean and division with standard deviation of the lung model.
- Combined model: Using the lung model on absorption and dark-field and the raw values of the phase and filter images.
- Phase filters: With only the differential phase and filter images

The results for the training set are presented in Table 1. Moving right from the first column additional features are included. I.e. the highest scoring model for 'Raw values' is the one based on Drk, Abs and R(25). Abs stands for absorption, Drk for dark-field, R is range filter, S is standard deviation filter, and E is entropy filter; the number given in brackets () denotes the radius of the filter. Highest accuracy is marked with blue, and the model selected by BIC and AIC – which in all cases coincided - is marked with green.

B. Test set

The models selected by BIC and AIC were tested on the test set. Results are presented in Table 2. The best method, based on the use of a lung model, yielded an accuracy of 84.05 %. Using the raw signal values results in a somewhat worse classification. Surprisingly, combining the lung model with phase information actually performs worse, and using the phase information alone is only a little better than guessing. An example for the classification algorithm for an emphysematous and a fibrotic mouse that showed good agreement with the findings of a radiologist are shown in Figure 7 and 8 respectively.
The lung model performed particularly poorly for two mice – one emphysematous and one fibrotic shown in Figure 9. The emphysema mouse revealed significantly lower dark-field values than expected.

IV. Discussion

Using the raw values alone gives an acceptable classification in most cases, but incorporating regional information through a lung model yields superior results. The differential phase-contrast information did not improve classification.

The raw values give a decent accuracy of 76.42 %, with a clear outlier in one fibrotic animal. Excluding that, the accuracy level increases to 80.15 %.

Using only absorption and dark-field combined with local information by terms of the lung model, has an overall accuracy of 84.05 %. In one emphysema mouse the model fails due to very high values in the dark-field. The reason for this could be additional lung injury (e.g. blood accumulation) in the lung. The one misclassified fibrotic animal (Figure 9) has a very early stage of fibrosis that looks akin to emphysema. Both fibrosis and emphysema start by an acute inflammation of the lung, therefore, classification of the two disorders at an early stage without prior clinical knowledge is challenging.

Incorporating the phase features does not increase the accuracy nor the robustness of the methods. This is surprising as there is discriminatory power in the features when inspecting the histograms. It is however likely that the variance from mouse to mouse is too large, making the phase features less useful. Using a more sophisticated classification method than simply multiplying the histogram values it might be possible to use this information. To test if the joint probabilities yielded better results, we tried to classify using a Quadratic Discriminant Analysis [20]. This did not yield better results.

A big advantage of our classification method is the associated probability of the classification and the easy to interpret visual results.

The phase features might have other applications as well. In [21] the aeration of the lungs was correlated with the power spectrum. A more direct measure could have been one of the phase features we employed, which would have yielded information of the local aeration.

V. Conclusion

Healthy, emphysematous, and fibrotic mice can be discriminated and classified using absorption and dark-field images. The classification is improved by incorporating regional information through a lung model, and can be presented as easy to interpret images. The differential phase-contrast information was useless in the classification, but might be useful if a more advanced classification scheme is employed.

References

[1] B. Zvezdin, S. Milutinov, M. Kojicic, M. Hadnadjev, S. Hromis, M. Markovic, and O. Gajic, "A postmortem analysis of major causes of early death in patients hospitalized with COPD exacerbation," *Chest*, vol. 136, no. 2, pp. 376–380, 2009.

[2] B. R. Celli, W. MacNee, a. Agusti, a. Anzueto, B. Berg, a. S. Buist, P. M. a Calverley, N. Chavannes, T. Dillard, B. Fahy, a. Fein, J. Heffner, S. Lareau, P. Meek, F. Martinez, W. McNicholas, J. Muris, E. Austegard, R. Pauwels, S. Rennard, a. Rossi, N. Siafakas, B. Tiep, J. Vestbo, E. Wouters, and R. ZuWallack, "Standards for the diagnosis and treatment of patients with COPD: A summary of the ATS/ERS position paper," *Eur. Respir. J.*, vol. 23, no. 6, pp. 932–946, 2004.

[3] R. W. George, "Diagnostic Imaging in COPD," *Changes*, vol. 29, no. 6, pp. 997–1003, 2012.

[4] T. E. E. Al King, "American Thoracic Societv Idiopathic Pulmonary Fibrosis : Diagnosis and Treatment," *Proc. Am. Thorac. Soc.*, vol. 161, pp. 646–664, 2000.

[5] A. Devaraj, "Imaging: How to recognise idiopathic pulmonary fibrosis," *Eur. Respir. Rev.*, vol. 23, no. 132, pp. 215–219, 2014.

[6] G. Raghu, H. R. Collard, J. J. Egan, F. J. Martinez, J. Behr, K. K. Brown, T. V. Colby, J. F. Cordier, K. R. Flaherty, J. a. Lasky, D. a. Lynch, J. H. Ryu, J. J. Swigris, A. U. Wells, J. Ancochea, D. Bouros, C. Carvalho, U. Costabel, M. Ebina, D. M. Hansell, T. Johkoh, D. S. Kim, T. E. King, Y. Kondoh, J. Myers, N. L. Müller, A. G. Nicholson, L. Richeldi, M. Selman, R. F. Dudden, B. S. Griss,

S. L. Protzko, and H. J. Schünemann, "An Official ATS/ERS/JRS/ALAT Statement: Idiopathic pulmonary fibrosis: Evidence-based guidelines for diagnosis and management," *Am. J. Respir. Crit. Care Med.*, vol. 183, no. 6, pp. 788–824, 2011.

[7] A. Yaroshenko, F. G. Meinel, M. Bech, A. Tapfer, A. Velroyen, S. Schleede, A. Bohla, A. Ö. Yildirim, K. Nikolaou, O. Eickelberg, and M. F. Reiser, "Pulmonary emphysema Diagnosis with a Preclinical," vol. 269, no. 2, pp. 427–433, 2013.

[8] H. Einarsdottir, A. Yaroshenko, M. Bech, K. Hellbach, G. Felix, O. Eickelberg, M. Reiser, R. Larsen, and F. Pfeiffer, "A Computer-Aided Diagnosis Scheme of Pulmonary Diseases from Grating Based Radiography," *Phys. Med. Biol.*, vol. In reveiw, 2015.

[9] K. Hellbach, A. Yaroshenko, F. G. Meinel, A. Ö. Yildirim, T. M. Conlon, M. Bech, M. Mueller, A. Velroyen, M. Notohamiprodjo, F. Bamberg, S. Auweter, M. Reiser, O. Eickelberg, and F. Pfeiffer, "In Vivo Dark-Field Radiography for Early Diagnosis and Staging of Pulmonary Emphysema," *Invest. Radiol.*, vol. 50, no. 7, pp. 430–435, Jul. 2015.

[10] F. Pfeiffer, T. Weitkamp, O. Bunk, and C. David, "Phase retrieval and differential phase-contrast imaging with low-brilliance X-ray sources," *Nat. Phys.*, vol. 2, no. 4, pp. 258–261, 2006.

[11] F. Pfeiffer, M. Bech, O. Bunk, P. Kraft, E. F. Eikenberry, C. Brönnimann, C. Grünzweig, and C. David, "Hard-X-ray dark-field imaging using a grating interferometer.," *Nat. Mater.*, vol. 7, no. 2, pp. 134–137, 2008.

[12] F. G. Meinel, a Yaroshenko, K. Hellbach, M. Bech, M. Müller, a Velroyen, F. Bamberg, O. Eickelberg, K. Nikolaou, M. F. Reiser, F. Pfeiffer, and a O. Yildirim, "Improved Diagnosis of Pulmonary Emphysema Using In Vivo Dark-Field Radiography.," *Invest. Radiol.*, vol. 00, no. 00, pp. 1–6, 2014.

[13] A. Tapfer, M. Bech, B. Pauwels, X. Liu, P. Bruyndonckx, A. Sasov, J. Kenntner, J. Mohr, M. Walter, J. Schulz, and F. Pfeiffer, "Development of a prototype gantry system for preclinical x-ray phase-contrast computed tomography," *Med. Phys.*, vol. 38, no. 11, p. 5910, 2011.

[14] A. Tapfer, M. Bech, A. Velroyen, J. Meiser, J. Mohr, M. Walter, J. Schulz, B. Pauwels, P. Bruyndonckx, X. Liu, A. Sasov, and F. Pfeiffer, "Experimental results from a preclinical X-ray phase-contrast CT scanner.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 109, no. 39, pp. 15691–6, 2012.

[15] F. G. Meinel, A. Yaroshenko, K. Hellbach, M. Bech, M. Müller, A. Velroyen, F. Bamberg, O. Eickelberg, K. Nikolaou, M. F. Reiser, F. Pfeiffer, and A. Ö. Yildirim, "Improved Diagnosis of Pulmonary Emphysema Using In Vivo Dark-Field Radiography," *Invest. Radiol.*, vol. 49, no. 10, pp. 653–658, Oct. 2014.

[16] A. Velroyen, A. Yaroshenko, D. Hahn, A. Fehringer, A. Tapfer, M. Müller, P. B. Noël, B. Pauwels, A. Sasov, A. Ö. Yildirim, O. Eickelberg, K. Hellbach, S. D. Auweter, F. G. Meinel, M. F. Reiser, M. Bech, and F. Pfeiffer, "Grating-based X-ray Dark-field Computed Tomography of Living Mice," *EBioMedicine*, 2015.

[17] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max- flow algorithms for energy minimization in vision," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 9, pp. 1124–1137, 2004.

[18] G. Schwarz, "Estimating th dimension of a model," *Ann. Stat.*, vol. 6, no. 2, pp. 461–464, 1978.

[19] C. M. Hurvich and C. L. Tsai, "Bias of the corrected aic criterion for underfitted regression and time series models," *Biometrika*, vol. 78, no. 3, pp. 499–509, 1991.

[20] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2. ed. 2009.

[21] S. B. Hooper, M. J. Kitchen, M. J. Wallace, N. Yagi, K. Uesugi, M. J. Morgan, C. Hall, K. K. W. Siu, I. M. Williams, M. Siew, S. C. Irvine, K. Pavlov, and R. a. Lewis, "Imaging lung aeration and lung liquid clearance at birth," *FASEB J.*, vol. 21, no. 12, pp. 3329–3337, 2007.

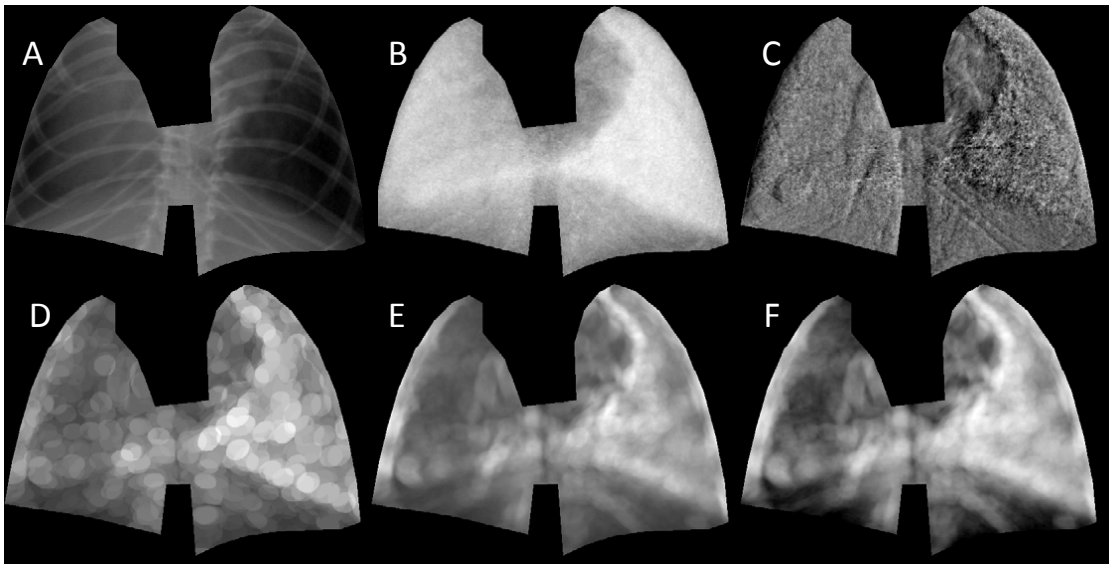


Figure 4: Segmented lungs from healthy mouse. A) absorption, B) dark-field, C) differential phase, D) range filter, E) STD filter, F) entropy filter. All filters with radius 15 pixels.

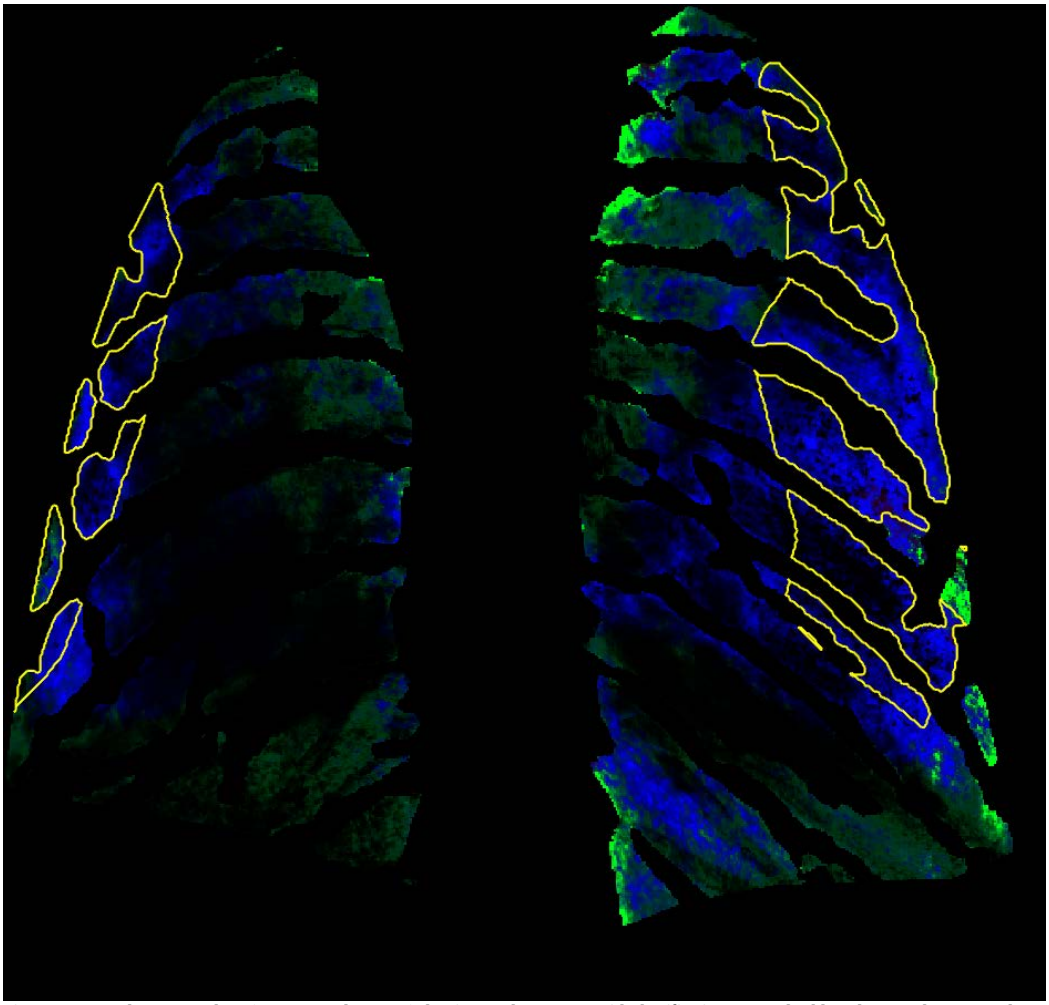


Figure 7: An emphysematous lung (mouse: Emphysema 4) showing good agreement with classification. Green: healthy. Blue: Emphysema. Red: fibrosis. The yellow line is the region demarcated as fibrosis by an expert.

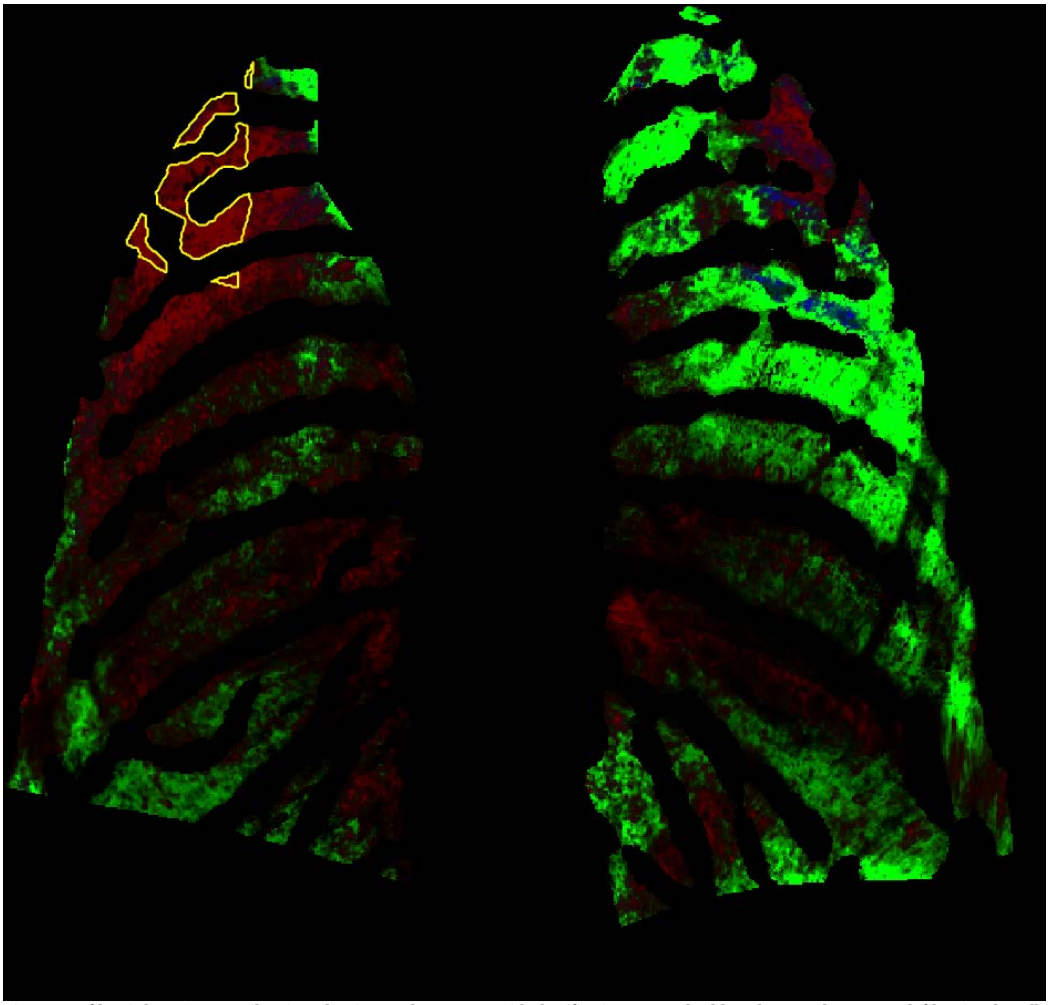


Figure 8: A fibrotic lung (mouse: Fibrosis 4) showing good agreement with classification. Green: healthy. Blue: Emphysema. Red: fibrosis. The yellow line is the region demarcated as fibrosis by an expert.

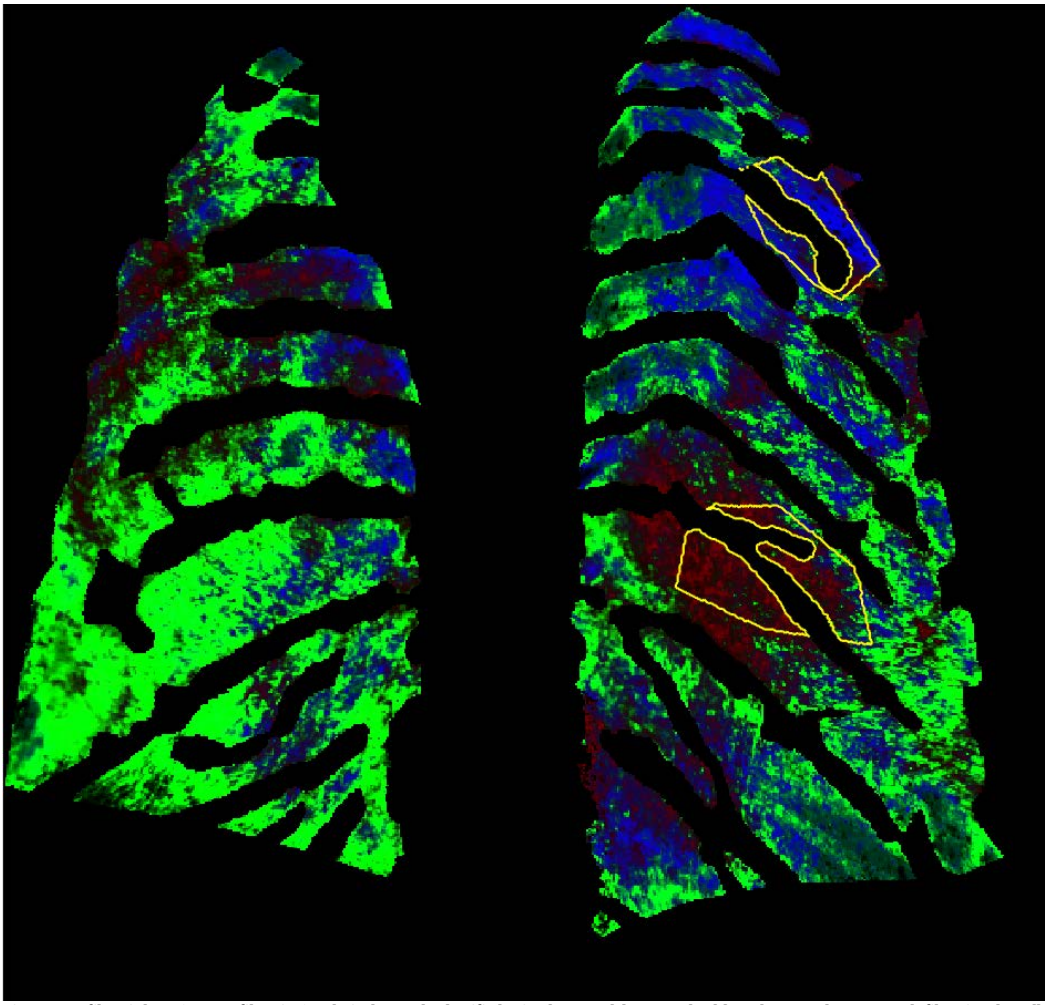


Figure 9: A fibrosis lung (mouse: fibrosis 1), relatively poorly classified using lung model. Green: healthy. Blue: Emphysema. Red: fibrosis. The yellow line is the region demarcated as fibrosis by an expert.

Table 1. Training set results

Raw values	Drk	Abs	R(25)	Phase	E(3)	R (20)	R(3)	S(25)	S(3)	R(22)
Accuracy %	66.91	79.57	81.24	81.22	80.23	78.98	77.77	76.75	75.64	74.66
Lung model	Drk	Abs	Phase	R(25)	R(24)	R(23)	R(22)	R(21)	R(20)	R(19)
Accuracy %	78.38	88.63	88.77	66.66	63.40	61.37	59.81	58.51	57.46	56.53
Combined	Drk	Abs	R(25)	Phase	S(25)	Ent(3)	R(3)	E(25)	S(3)	R(20)
Accuracy %	78.37	88.63	90.63	90.71	90.48	89.95	88.78	87.04	85.26	83.30
Phase filters	R(25)	Phase	R(23)	R(24)	R(3)	S(25)	R(20)	R(22)	S(3)	R(21)
Accuracy %	61.05	62.26	62.56	62.27	62.12	62.14	62.01	61.83	61.68	61.56

Table 2: Accuracy values obtained for test set

Disease Labelled Mice	Raw [Drk Abs]	Lung model [Drk Abs]	Combined [Drk Abs R (25)]	Phase measures [R (25)]
Healthy 1	75.66	94.10	93.15	28.63
Healthy 2	87.13	94.53	93.66	33.01
Healthy 3	50.39	76.70	76.98	28.80
Healthy 4	48.99	85.61	82.41	27.45
Healthy 5	82.96	98.26	97.34	30.65
Healthy 6	64.47	96.52	95.15	27.98
Emphysema 1	84.33	82.15	67.78	27.54
Emphysema 2	96.49	95.96	97.63	83.63
Emphysema 3	91.83	92.95	36.18	0.47
Emphysema 4	99.59	94.60	84.99	20.19
Emphysema 5	80.66	65.80	44.31	9.80
Emphysema 6	79.36	33.89	57.03	66.49
Emphysema 7	95.01	84.80	62.29	9.57
Emphysema 8	84.99	58.22	80.56	83.51
Fibrosis 1	1.84	51.67	27.94	2.50
Fibrosis 2	98.59	84.98	91.71	90.97
Fibrosis 3	60.94	96.04	98.96	97.11
Fibrosis 4	98.04	98.93	23.69	22.96
Fibrosis 5	82.51	86.59	96	82.54
Fibrosis 6	41.00	92.84	66.47	1.08
Fibrosis 7	100	99.96	99.96	82.02
All Healthy	68.27	90.953	89.78	29.42
All Emphysemous	89.03	76.05	66.35	37.65
All Fibrotic	68.99	87.29	72.10	54.17
Total	76.42	84.05	74.96	40.80

Bibliography

[Akaike, 1974] Akaike, H. (1974). 'A new look at the statistical model identification'. *IEEE Transactionson Automatic Control*, 19(6):716–723.

[Anderson, 1962] Anderson, T. W. (1962). On the Distribution of the Two-Sample Cramer-von Mises Criterion. *The Annals of Mathematical Statistics*, 33(3):1148–1159.

[Anderson and Darling, 1952] Anderson, T. W. and Darling, D. A. (1952). Asymptotic Theory of Certain "Goodness of Fit" Criteria Based on Stochastic Processes. *The Annals of Mathematical Statistics*, 23(2):193–212.

[Bech, 2009] Bech, M. (2009). *X-ray imaging with a grating interferometer*. PhD thesis.

[Boellaard et al., 2001] Boellaard, R., van Lingen, a., and Lammertsma, a. a. (2001). Experimental and clinical evaluation of iterative reconstruction (OSEM) in dynamic PET: quantitative characteristics and effects on kinetic modeling. *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*, 42(5):808–817.

[Box, 2015] Box, G. E. P. (2015). A General Distribution Theory for a Class of Likelihood Criteria. 36(3):317–346.

[Box and Cox, 1964] Box, G. E. P. and Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 26(2):211–252.

[Boykov and Kolmogorov, 2003] Boykov, Y. and Kolmogorov, V. (2003). Computing geodesics and minimal surfaces via graph cuts. *Proceedings Ninth IEEE International Conference on Computer Vision*, 2003.

[Boykov and Kolmogorov, 2004] Boykov, Y. and Kolmogorov, V. (2004). An experimental comparison of min-cut/max- flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(9):1124–1137.

[Brown and Forsythe, 1974] Brown, M. B. and Forsythe, a. B. (1974). Robust Tests for Equality of Variances. *Journal of the American Statistical Association*, 69(346):364–367.

[Burnham, 2004] Burnham, K. P. (2004). Multimodel Inference: Understanding AIC and BIC in Model Selection. *Sociological Methods & Research*, 33(2):261–304.

[Bushberg et al., 2011] Bushberg, J., Seibert, J., Leidholdt, E., and Boone, J. (2011). *The Essential Physics of Medical Imaging*. 3. edition.

[Chan and Vese, 2001] Chan, T. F. and Vese, L. a. (2001). Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277.

[Clemmensen et al., 2011] Clemmensen, L., Hastie, T., Witten, D., and Ersbøll, B. (2011). Sparse Discriminant Analysis. *Technometrics*, 53(4):406–413.

[Comtat et al., 2008] Comtat, C., Sureau, F. C., Sibomana, M., Hong, I. K., Sjoholm, N., and Trebossen, R. (2008). Image based resolution modeling for the HRRT OSEM reconstructions software. *2008 IEEE Nuclear Science Symposium Conference Record*, (5):4120–4123.

[Cootes and Taylor, 2001] Cootes, T. and Taylor, C. (2001). Statistical models of appearance for medical image analysis and computer vision. *Medical Imaging 2001*, pages 236–248.

[Cootes et al., 1995] Cootes, T., Taylor, C., Cooper, D., and Graham, J. (1995). Active Shape Models-Their Training and Application.

[Cootes et al., 1998] Cootes, T. F., Edwards, G. J., and Taylor, C. J. (1998). Active appearance models. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 23, pages 484–498.

[Diven and Almy, 1950] Diven, B. C. and Almy, G. M. (1950). Photodisintegration of silver and aluminum. *Physical Review*, 80(3):407–414.

[Dixon, 1984] Dixon, W. T. (1984). Simple proton spectroscopic imaging. *Radiology*, 153(1):189–194.

[Driscoll and Borror, 2000] Driscoll, M. F. and Borror, C. M. (2000). Sums of squares and expected mean squares in SAS. *Quality and Reliability Engineering International*, 16(5):423–433.

[Ebastien, 1999] Ebastien, S. (1999). Stereo Without Epipolar Lines : A Maximum-Flow Formulation. *Framework*, 34(2/3):147–161.

[Einarsdottir et al., 2015] Einarsdottir, H., Yaroshenko, A., Bech, M., Hellbach, K., Felix, G., Eickelberg, O., Reiser, M., Larsen, R., and Pfeiffer, F. (2015). A Computer-Aided Diagnosis Scheme of Pulmonary Diseases from Grating Based Radiography. *Physics in Medicine and Biology*, In reveiw.

[Elliott, 2011] Elliott, a. (2011). *Committee on Medical Aspects of Radiation in the Environment: Fourteenth Report.*

[Ersbøll and Conradsen, 2012] Ersbøll, B. and Conradsen, K. (2012). *Multivariate statistics. An introduction.* Technical University of Denmark, Department of Applied Mathematics and Computer Science, 8 edition.

[Fawcett, 2006] Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8):861–874.

[Feldkamp et al., 1984] Feldkamp, L. a., Davis, L. C., and Kress, J. W. (1984). Practical cone-beam algorithm. *Journal of the Optical Society of America A*, 1(6):612.

[Fischl et al., 2002] Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., Van Der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., and Dale, A. M. (2002). Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron*, 33(3):341–355.

[Fuller, 1985] Fuller, E. (1985). Photonuclear reaction cross sections for 12C, 14N and 16O.

[Geman and Geman, 1984] Geman, S. and Geman, D. (1984). Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-6(6):721–741.

[Greig et al., 1989] Greig, D. M., Porteous, B. T., and Seheult, A. H. (1989). Exact Maximum A Posteriori Estimation for Binary Images. *Journal of the Royal Statistical Society*, 51(2):271–279.

[Hansen et al., 2008] Hansen, A. T., Hansen, S. B., and Petersen, J. B. (2008). The potential application of silver and positron emission tomography for in vivo dosimetry during radiotherapy. *Physics in medicine and biology*, 53(2):353–60.

[Hanson, 2009] Hanson, L. G. (2009). Introduction to Magnetic Resonance Imaging Techniques, Educational material. *Http://Www.Drcmr.Dk:8080/37*, pages 1–48.

[Hastie et al., 2009] Hastie, T., Tibshirani, R., and Friedman, J. (2009). The Elements of Statistical Learning. *Elements*, 1:337–387.

[Hudson and Larkin, 1994] Hudson, H. M. and Larkin, R. S. (1994). Accelerated image reconstruction using ordered subsets of projection data. *IEEE transactions on medical imaging*, 13(4):601–9.

[Hurvich and Tsai, 1991] Hurvich, C. M. and Tsai, C. L. (1991). Bias of the corrected aic criterion for underfitted regression and time series models. *Biometrika*, 78(3):499–509.

[Iacobucci, 1995] Iacobucci, D. (1995). Analysis of Variance for Unbalanced Data.

[Ishikawa, 2003] Ishikawa, H. (2003). Exact optimization for Markov Random Fields with convex priors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(10):1333–1336.

[Jaffray et al., 2002] Jaffray, D. a., Siewerdsen, J. H., Wong, J. W., and Martinez, A. a. (2002). Flat-panel cone-beam computed tomography for image-guided radiation therapy. *International Journal of Radiation Oncology Biology Physics*, 53(5):1337–1349.

[Jensen, 2010] Jensen, T. (2010). *Refraction and scattering based x-ray imaging*. PhD thesis.

[Johnson et al., 2007] Johnson, R. A., Wichern, D. W., and Hall, P. P. (2007). *Applied Multivariate Statistical Analysis*.

[Johnson and Gibbons, 1973] Johnson, W. D. and Gibbons, J. D. (1973). *Nonparametric Statistical Inference*, volume 15.

[Jølck et al., 2014] Jølck, R. I., Binderup, T., Hansen, A. E., Scherman, J. B., Munch, P., Kjæ r, A., and Andresen, T. L. (2014). Injectable Colloidal Gold in a Sucrose Acetate Isobutyrate Gelating Matrix with Potential Use in Radiation Therapy. pages 1680–1687.

[Jørgensen, 2005] Jørgensen, B. (2005). *CT-teknik*. 7 edition.

[Kinahan et al., 2003] Kinahan, P. E., Hasegawa, B. H., and Beyer, T. (2003). X-ray-based attenuation correction for positron emission tomography/computed tomography scanners. *Seminars in nuclear medicine*, 33(3):166–79.

[Knöß, 2004] Knöß, C. (2004). *Evaluation and Optimization of the High Resolution Research Tomograph (HRRT)*. PhD thesis.

[Labatut and Cherifi, 2011] Labatut, V. and Cherifi, H. (2011). Accuracy measures for the comparison of classifiers. *The 5th International Conference on Information Technology*, (May):11.

[Larsen, 2015] Larsen, C. T. (2015). *Development and Application of Tools for MRI Analysis*. PhD thesis.

[Li et al., 2006] Li, K., Wu, X., Chen, D. Z., and Sonka, M. (2006). Optimal surface segmentation in volumetric images - A graph-theoretic approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(1):119–134.

[Lilliefors, 1967] Lilliefors, H. W. (1967). on the Kolmogorov-Smirnov Test for Normality With Mean and Variance Unknown. *Journal of the American Statistical Association*, 62(318):399–402.

[Liu et al., 2007] Liu, C., Frazier, P., and Kumar, L. (2007). Comparative assessment of the measures of thematic classification accuracy. *Remote Sensing of Environment*, 107(4):606–616.

[Ma, 2008] Ma, J. (2008). Dixon techniques for water and fat imaging. *Journal of Magnetic Resonance Imaging*, 28(3):543–558.

[Matthews, 1975] Matthews, B. W. (1975). Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochimica et biophysica acta*, 405(2):442–451.

[Nesterets and Wilkins, 2008] Nesterets, Y. I. and Wilkins, S. W. (2008). Phase-contrast imaging using a scanning-double-grating configuration. *Optics express*, 16(8):5849–5867.

[Nishimura, 2010] Nishimura, D. (2010). *Principles of Magnetic Resonance Imaging*.

[Oja, 2010] Oja, H. (2010). *Multivariate Nonparametric Methods with R*, volume 199 of *Lecture Notes in Statistics*. Springer New York, New York, NY.

[Oja and Randles, 2004] Oja, H. and Randles, R. H. (2004). Multivariate Nonparametric Tests. *Statistical Science*, 19(4):598–605.

[Pitman, 1993] Pitman, J. (1993). *Probability*. Springer New York, New York, NY.

[Reuter et al., 2012] Reuter, M., Schmansky, N. J., Rosas, H. D., and Fischl, B. (2012). Within-subject template estimation for unbiased longitudinal image analysis. *NeuroImage*, 61(4):1402–1418.

[SAS, 1999] SAS (1999). *The MIXED procedure*.

[Sas, 2007] Sas (2007). *The CALIS procedure*.

[SAS, 2008] SAS (2008). The GLM Procedure.

[Schulze et al., 2011] Schulze, R., Heil, U., Groß, D., Bruellmann, D. D., Dranischnikow, E., Schwanecke, U., and Schoemer, E. (2011). Artefacts in CBCT: A review. *Dentomaxillofacial Radiology*, 40(5):265–273.

[Schwarz, 1978] Schwarz, G. (1978). Estimating th dimension of a model. *The annals of Statistics*, 6(2):461–464.

[Shapiro and Wilk, 1965] Shapiro, S. S. and Wilk, M. B. (1965). An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, 52(3/4):591–611.

[Siemens, 2009] Siemens (2009). Inside Biograph TruePoint PET • CT.

[van Velden et al., 2009] van Velden, F. H. P., Kloet, R. W., van Berckel, B. N. M., Lammertsma, A. a., and Boellaard, R. (2009). Accuracy of 3-dimensional reconstruction algorithms for the high-resolution research tomograph. *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*, 50(1):72–80.

[Zuidervel, 1994] Zuidervel, K. (1994). Contrast limited adaptive histogram equalization. In *Graphics gems IV*, pages 474–485.